(Continued)

CONTENTS—*Continued from preceding page*

CONTENTS—*Continued from preceding page*

CONTENTS—*Continued from preceding page*

Editorial—*Journal* Distribution Change

Starting with Volume 99 in 1996, Society members (other than students) were given three optional ways to receive the *Journal*: (a) continue receiving the entire *Journal* in monthly issues in printed form; (b) receive the entire *Journal* in cumulative bimonthly CD-ROM issues, the sixth issue containing both volumes for that calendar year; (c) receive monthly an offprint of the front part of the *Journal* containing acoustical news (USA, International, Standards), views (FORUM section), reviews (of books and patents), tutorial or research articles selected for broad interest, and the table of contents for the entire issue of the *Journal*. Each option had different advantages and was chosen by a significant percentage of the membership.

Combinations of options were also offered but were chosen initially by relatively few members. CD-ROM recipients (including all student members) heard from the Society less frequently, and required access to computers for reading and printout. A suggestion was received from numerous individuals for the *printed* offprint portion to be available to all members including those members and student members now receiving it only on the CD-ROM.

After a mail questionnaire sampling survey of the membership the Executive Council decided to adopt this suggestion, starting with Volume 103 in January 1998. The offprint portion of the *Journal* may be renamed to distinguish it by content rather than by the process of its production, but it will still be the front part of the *Journal.* It will continue to evolve in format and content for the benefit of the members and the Society.

DANIEL W. MARTIN
Editor-in-Chief

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.*

Preliminary notice: 134th meeting of the Acoustical Society of America

The 134th meeting of the Acoustical Society of America (ASA) will be held Monday through Friday, 1–5 December 1997 at the Town and Country Hotel, San Diego, California. Registration will begin Sunday, 30 November, at 3:00 p.m.

For information about the technical program refer to the ''Meetings'' section on the ASA Home Page at http://asa.aip.org/

Technical program and special sessions

The technical program will consist of lecture and poster sessions. Technical sessions will be scheduled Monday through Friday, 1–5 December. The following special sessions are planned for invited and contributed papers:

Special Sessions

Subjects	Sponsoring Technical Committees/Groups
Acoustic observations of ocean ridge processes	Acoustical Oceanography
Ambient noise inversions	Acoustical Oceanography
Stochastic inverse methods applied to ocean processes	Acoustical Oceanography/ Underwater Acoustics
Acoustics of reptiles and amphibians	Animal Bioacoustics
Biologically inspired acoustics models and systems	Animal Bioacoustics
Instrumentation in animal bioacoustics	Animal Bioacoustics
C. Scott Johnson session: Whale and dolphin acoustics	Animal Bioacoustics
Recording and production studios— poster session	Architectural Acoustics
Paul Veneklassen memorial session	Architectural Acoustics
Loudspeakers for listening spaces	Architectural Acoustics/ Engineering Acoustics
Impact of real time signal processing on acoustical measurement techniques	Architectural Acoustics/Signal Processing in Acoustics
Hands on acoustics for high school students	Education in Acoustics
Informal science education in acoustics	Education in Acoustics
Take fives: Sharing ideas for teaching acoustics	Education in Acoustics
Acoustic systems designed for harsh environments	Engineering Acoustics
Advanced techniques in ultrasonic nondestructive evaluation	Engineering Acoustics
Advances in microfabricated accelerometers	Engineering Acoustics
Fiber optic sensors for acoustic applications	Engineering Acoustics
Computer jazz improvisation	Musical Acoustics
Performance anxiety	Musical Acoustics
Piano	Musical Acoustics
Allocation of hearing loss—Where does hearing loss come from?	Noise
Annoyance from construction noise	Noise
Computational aeroacoustics	Noise/Engineering Acoustics
Measurement errors and calibration tolerances	Noise/Engineering Acoustics
Session in honor of Robert W. Young	Noise/Musical Acoustics/ Underwater Acoustics
Effects of fetal noise and vibration exposure	Noise/Psychological and Physiological Acoustics
Acoustics in multimedia— Perceptual issues	Signal Processing in Acoustics/Psychological and Physiological Acoustics/ Speech Communication
Signal processing for multi-channel vibrational analysis	Signal Processing in Acoustics/Structural Acoustics and Vibration
Basic science at the intersection of speech science and communication disorders—workshop	Speech Communication
Phonetic perception and word recognition	Speech Communication
Sources of individual variability in speech production and perception	Speech Communication
Computational vibroacoustics (FEM, BEM, SEA)	Structural Acoustics and Vibration
Direct and inverse methods in strongly range-dependent environments	Underwater Acoustics/ Acoustical Oceanography
Advances in underwater acoustic range technology: Communication, tracking and installation	Underwater Acoustics/ Engineering Acoustics

Other lectures

The Technical Committee on Architectural Acoustics will sponsor a Vern O. Knudsen Distinguished Lecture to be presented by Christopher Jaffe of Jaffe Acoustics, Norwalk, Connecticut.

Tutorial lecture, short courses, hot topics

The Tutorials Committee has planned a Tutorial Lecture on ''The Dolphin Echolocation System'' and three short courses: (1) Inverse Methods in Ocean Acoustics/Acoustical Oceanography; (2) Array Measurements and Near-field Acoustical Holography; and (3) Basic Digital Signal Processing in Acoustics. See the following articles in this news section for full details on these lectures and short courses. A ''Hot Topics'' session sponsored by the Tutorials Committee is also planned.

Program

An advance meeting program summary will be published in the October issue of JASA and a complete meeting program will be mailed as Part 2 of the November issue. Abstracts will be available on the ASA Home Page (http://asa.aip.org) in mid-October.

Student transportation subsidies

A student transportation subsidies fund has been established to provide limited funds to students to partially defray transportation expenses to meetings. Students presenting papers who propose to travel in groups using economical ground transportation will be given first priority to receive subsidies, although these conditions are not mandatory. No reimbursement is intended for the cost of food or housing. The amount granted each student depends on the number of requests received. To apply for a subsidy, submit a written proposal to be received by 15 September: Elaine Moran, ASA, 500 Sunnyside Blvd., Woodbury, NY 11797, 516-576-2360, FAX: 516-576-2377, E-mail: asa@aip.org. The proposal should indicate your status as a student, whether you have submitted an abstract, whether you are a member of ASA, method of travel, whether you will travel alone or with other students, names of those traveling with you, and approximate cost of transportation.

Students meet members for lunch

The Education Committee has established a program for students to meet with members of the ASA over lunch. Students should sign up at the registration desk at the meeting. Members who wish to participate should contact Scott D. Sommerfeldt, Tel.: 801-378-2205, E-mail: s_sommerfeldt@byu.edu prior to the meeting. Participants are responsible for cost of their own meal.

Plenary session, awards, Fellows' luncheon and social events

Complimentary buffet socials with cash bar will be held early on Tuesday and Thursday evenings, 2 and 4 December. The Plenary session will be held on Wednesday afternoon, 3 December, where Society awards will be presented and recognition of Fellows will be announced. A Fellows' Luncheon will be held on Thursday, 4 December; ASA Fellows may purchase tickets at the meeting.

Exhibits

An equipment exhibition for the meeting will be held at the Town and Country Hotel, 2–3 December. The exhibition will include active noise control systems, computer-based instrumentation, sound level meters, sound intensity systems, signal processing systems, devices for noise control and acoustical materials. Organizations interested in exhibiting should contact: Robert Finnegan, Advertising and Exhibits, American Institute of Physics, 500 Sunnyside Blvd., Woodbury, NY 11797; Tel. 516-576-2433; FAX 516-576-2481; rgf@aip.org.

Paper copying service

Authors are requested to provide one paper copy of their projection material and/or paper(s) to the Paper Copies Desk in the Registration area upon arrival. The copy should be material on one side only on 8 1/2×11 inch or A4 paper suitable for photocopy reproduction. Copies of available papers will be made for a nominal charge.

Technical tours

Technical tours have been arranged for ANOMS (Airport Noise and Operations Monitoring System); Hubbs-Sea World Research Institute, and to the Navy Marine Mammal Research Laboratory. A schedule of tours and sign-up sheets will be available at the meeting.

Registration

Registration will begin Sunday afternoon, 30 November, at 3:00 p.m. at the Town and Country Hotel. Checks or travelers checks in U.S. funds drawn on U.S. banks and Visa, MasterCard, and American Express credit cards will be accepted for payment of registration. The registration fee is $200 for members of the Acoustical Society of America. The registration fee is $245 for nonmembers and $35 for Emeritus members of ASA (Emeritus status pre-approved by ASA) and accompanying persons. One-day registration is available at $100 for members and $125 for nonmembers. A nonmember who pays the $245 nonmember registration fee and simultaneously applies for Associate Membership in the Acoustical Society of America will be exempt from dues payment for the first year of membership (1998). Invited speakers who are members of the Acoustical Society of America are expected to pay the registration fee, but nonmember invited speakers who participate in the meeting for one day only may register without charge. Nonmember invited speakers who wish to participate in the meeting for more than one day will be charged the member registration fee of $200, which will include a one-year membership in the ASA upon completion of an application form at the meeting. There is no fee for students with active student identification cards.

Assistive listening devices

Anyone planning to attend the meeting who will require the use of an assistive listening device, is requested to advise the Society in advance of the meeting by writing to: Acoustical Society of America, 500 Sunnyside Boulevard, Woodbury, NY 11797.

Accompanying persons program

Accompanying persons are welcome. A Hospitality Room will be open from 4:00 to 6:00 p.m. Sunday, 30 November and daily Monday through Friday, 8:00 to 11:00 a.m. Special activities will be planned as well as tours to Wild Animal Park, San Diego Zoo, Sea World, Tijuana, Old Town, and the Aquarium.

Air transportation

The Lindbergh International Airport is served by all of the major airlines.

Ground transportation

The Town and Country Hotel is located at the intersection of Interstate 8 and Interstate 5. Transportation from the airport to the hotel may be by taxi or by Cloud 9 limousine. The taxi fare is about $15 and the limousine is $7.

Hotel accommodations

The Town and Country Hotel is the Conference hotel and all of the rooms and suites have been reserved at the following special rates. Garden: $85/single or double; Towers: $95/single or double. (Children under 18 stay free when staying with parents and utilizing existing beds). To reserve a room call 1-800-772-8527 (619-291-7131; FAX: 619-291-3584) or send reservation by mail to: Town & Country Hotel, 500 Hotel Circle North, San Diego, CA 92108. The hotel accepts Visa, MasterCard, American Express, Discover, Diners Club, and Carte Blanche credit cards.

When making reservations please ensure that the Acoustical Society name is used so that you will be given the discounted convention rate. **Make reservations before the ASA room block cut-off date of 8 November**.

Weather

The weather in San Diego the first week of December (as always) is mild with very little humidity. The day temperature will range between 18–21 °C (65–70 °F) while the night temperature will hover around 13 °C (55 °F). Occasional desert winds (the infamous Santa Annas) that last 1–3 days may also occur during that time with the daytime temperature ranging to 27 °C (80 °F). The daily probability of rain (more than 0.5 cm) is 9%.

CHARLES E. SCHMID
Executive Director

Tutorial lecture on Dolphin Echolocation System

A Tutorial Lecture on The Dolphin Echolocation System will be presented by Whitlow W.L. Au, Hawaii Institute of Marine Biology, Kailua, Hawaii, at 7:00 p.m. on Monday, 1 December.

ABSTRACT

The dolphin echolocation system. Whitlow W.L. Au (Marine Mammal Research Program, Hawaii Inst. of Marine Biology, P.O. Box 1106, Kailua, HI 996734)

The sonar of dolphins may be the most sophisticated of all sonar systems, biological or man-made, in shallow waters and for short ranges. The Atlantic bottlenose dolphin emit short duration (50–70 μs), high frequency (120–140 kHz), broadband (40–50 kHz) echolocation signals with peak-to-peak amplitudes up to 228 dB *re*:1 μPa. The type of signals used by dolphins play a significant role in their sonar discrimination capabilities. They have been observed detecting, classifying, and retrieving prey that is buried in sandy bottom up to a depth of about 0.3 m. In addition, controlled echolocation experiments have shown that dolphins can discriminate wall thickness, material composition, shape, and size of targets. The echolocation system of dolphins will be discussed in three parts. The first part will consider the properties of the receiving system (auditory capabilities). The second part will deal with the characteristics of echolocation signals and the third part will consider several sonar capabilities.

Lecture notes

Notes summarizing the lecture will be available at the meeting. Those who register by 17 November are guaranteed receipt of a set of notes.

Tutorial lecture preregistration

To partially defray the cost of the lecture a registration fee is charged. The fee is $15 for registration received by 17 November and $25 thereafter including on-site registration at the meeting. The fee for students with current ID's is $7.00 for registration received by 17 November and $12.00 thereafter, including on-site registration at the meeting.

To register send your name, address, telephone, fax, and e-mail addresses, along with your payment (check or international money order in U.S. funds drawn on a U.S. bank or Visa, Mastercard, or American Express credit card) to: Short Course Registration, Acoustical Society of America, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org.

Short courses to be offered at Fall meeting in San Diego

Three short courses will be offered in conjunction with the Fall 1997 meeting to be held in San Diego, CA. The courses will be held on Friday and Saturday, 5 and 6 December. Information about each course and registration appears below.

Short Course on Basic Digital Signal Processing in Acoustics

Following the successful offering of a short course on basic digital signal processing in conjunction with the Spring 1995 Meeting, the Tutorials Committee has **tentatively** scheduled this repeat offering pending the availability of an instructor. This course will be canceled if there is insufficient registration.

This short course is intended primarily for acousticians whose work involves experimental data (i.e., signal) acquisition and digital processing to obtain useful results. While the course will include material not yet widely known, it will also include materials intended for individuals whose formal education did not adequately cover the subject or who feel a need for a refresher. Course emphasis will be on the discrete Fourier transform (DFT), its fast Fourier transform (FFT) cousin, and how they can be used to obtain useful estimates for signal and system properties. Signal design, signal analysis, and applications to system identification will be discussed. Both deterministic and random signals will be considered. Some prior acquaintance with linear algebra, particularly time-frequency transformations such as Fourier series and/or Fourier transforms, and representation by complex exponentials will be assumed.

For further information about the course and content contact: Dr. Joseph Pope, ASA Tutorials Committee; telephone: 617-969-2606 or FAX: 617-244-2606.

The registration fee for this course is $300 and covers attendance, instructional materials, and coffee breaks.

Short Course on Inverse Methods in Ocean Acoustics/Acoustical Oceanography

Inverse methods are widely used in the analysis of ocean acoustic data, including beamforming calculations, noise estimation and other signal processing, one-, two-, and three-dimensional sound-speed estimation, and estimation of time-dependent ocean fields with dynamical constraints. Inverse methods allow us to choose a solution for problems which have no exact solution, usually due to incomplete and/or inconsistent data. The price we pay for this is that we must make assumptions in advance about the criteria for the choice. The course will try to provide a basic introduction to the philosophy and methods of inverse theory, emphasizing the commonality between different applications, and demonstrating the concepts with simple examples. The goal of the course is to give the participants the background to choose among the methods in the literature. We will look in detail at a few examples of inverse problems of current interest, such as sound-speed estimation from travel times or from full field inverses, to illustrate how the basic issues apply in practical cases. Participants can also suggest their own examples in advance, to help motivate the discussion. Notation and concepts from linear algebra are central to most references on inverse methods, especially least-squares methods, but can be a barrier to understanding. Although the course will emphasize principles which can be explained simply, most solution techniques use matrices, and some familiarity with matrix notation and manipulation will be helpful.

The course instructor will be Bruce Cornuelle, Associate Research Oceanographer at the University of California at San Diego's Scripps Institute of Oceanography. He has been working with inverse methods in ocean acoustics for more than 15 years. For more information about the course program contact: Bruce Cornuelle, bcornuelle@ucsd.edu or FAX: (619) 534-7132.

The registration fee is $250 and covers attendance, instructional materials, and coffee breaks.

Short Course on Array Measurements and Near-field Acoustical Holography

Near-field acoustical holography (NAH) has proven to be a very powerful tool in sound and noise control/analysis since its invention in 1980. Its ability to provide both vibration and radiation analysis and source localization from a single pressure array measurement, has made it a very convenient, versatile, and powerful tool. With microphone/hydrophone arrays becoming more common in practice, acousticians are well positioned to be able to use NAH to attack their particular problems, hopefully obtaining timely, economical solutions in both the research community and in industry. In order to use NAH, one needs to understand the basics of the theory, and how to implement this theory in the experimental world. To this end, this short course is aimed at providing the knowledge necessary for one to be able to implement NAH successfully for his/her own noise analysis/control problem.

The course instructor is Earl G. Williams of the Naval Research Laboratory, Washington, DC. He is co-inventor of NAH and has been developing and applying NAH to problems of Navy/DOD interest since 1982. An invited lecturer throughout the world, Dr. Williams has written dozens of papers on NAH and has just completed a book on the subject.

For further information about the course program contact: Earl Williams, Code 7137, Naval Research Laboratory, Washington DC 20375, email: williams@astro.nrl.navy.mil or telephone (202) 767-0330.

The registration fee is $250 and covers attendance, instructional materials, and coffee breaks.

Registration

The number of attendees will be limited in each short course, so please register early to avoid disappointment. Only those who have registered by 17 November will be guaranteed receipt of instructional materials to the course. There will be a $50 discount for reservations made prior to 27 October. Full refunds will be made for cancellations prior to 17 November. Any cancellation after 17 November will be charged a $25 processing fee.

To register send your name, address, telephone, fax, and e-mail addresses and name of course for which you are registering, along with your payment (check or international money order in U.S. funds drawn on a U.S. bank or Visa, Mastercard, or American Express credit card) to: Short Course Registration, Acoustical Society of America, 500 Sunnyside Blvd., Woodbury, NY 11797, Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org.

Job Opening

Soundwich, Inc.: Rapidly growing automotive supplier has opportunity for Acoustics and Vibration Engineer. Will set up equipment and perform acoustic and thermal testing to meet R&D needs. BSME degree with acoustics emphasis and experience in an acoustics/vibrations R&D environment. Please send resume and salary requirements to: Soundwich, Inc., Attn: Karen, 881 Wayside Rd., Cleveland, OH 44110.

Annual Reports of Technical Committees and Technical Groups

Acoustical Oceanography

The fall meeting at Honolulu was a major event for Acoustical Oceanography, with a greater than usual number of special sessions, some co-organized with colleagues from Japan. David Palmer and Iwao Nakano organized three sessions on the topic ''Acoustical Ocean Monitoring: Determination of Current and Temperature Fields.'' Ralph Stephen and Christopher Fox co-organized a joint UW–AO special session on ''Earthquake Acoustics,'' and Christian de Moustier and Hidekazu Tokuyama organized two AO sessions on ''Ocean Floor Surveys.'' James Lynch and Steven Stanic organized a joint UW–AO session, ''High-Frequency Propa-

gation and Inversion in Shallow Water,'' and Jules Jaffe worked with David Havelock to develop the joint AO–Signal Processing session ''Imaging the Ocean Volume.'' As part of a joint AO–UW technical initiative, Michael Longuet-Higgins delivered a Special Lecture on Acoustical Oceanography as part of two special sessions on ''Bubble Measurements Near the Ocean Surface,'' organized by David Farmer and Kerry Commander.

Permits for ocean acoustic research required by the Marine Mammal Protection Act have been a continuing concern of our technical committee. At the Honolulu meeting, Roger Gentry of the National Marine Fisheries Service presented a progress report on development of a streamlined permit process. Technical input will be required, and it is expected that ASA members will serve on NMFS advisory panels.

For the Penn State meeting, Jim Lynch and Mohsen Badiey organized two joint AO–UW sessions on ''Acoustical Measurement of Coastal Ocean Processes.'' Another significant event was Mike Buckingham's paper, ''Hot Topics in Acoustical Oceanography.'' The Penn Sate Technical Committee meeting was largely devoted to discussions of the programs of future meetings, particularly the Seattle and Berlin meetings, as the international character of these meetings provides special opportunities for AO. At the Penn State Technical Committee meeting, Jim Lynch reported on the AO home page and solicited further inputs and links to researcher's home pages.

Special thanks are due Art Newhall and Jim for doing an outstanding job on one of the first ASA technical committee home pages.

DARRELL JACKSON
Chair

Animal Bioacoustics

The Animal Bioacoustics technical committee conducted three special sessions at the joint ASA–ASJ meeting in Hawaii, in Dec. 1996. The sessions entitled ''Acoustics of Vocalizing Animals—I and —II'' were organized by Whitlow Au, from the Hawaii Institute of Marine Biology and Hiroshi Riquimaroux of Doshisha University. There was one invited speaker to the first session, and thirteen contributed papers. The second session included two invited papers and ten contributed papers. The third special session was entitled ''Auditory and Non-Auditory Effects'' and was organized by Ann Bowles from Hubbs Sea World and Kazuo Kanoya from Chiba University. There were two invited papers and eleven contributed papers. The total number of papers presented in the three Animal Bioacoustics sessions in Honolulu totaled forty five, the largest amount ever for Animal Bioacoustics.

The Animal Bioacoustics Technical Specialty Group gained official status as a technical committee following the Honolulu meeting. Whitlow Au was elected chairman of the AB technical committee in a mail-in election. Mardi Hastings was elected a Fellow of the ASA at the Honolulu meeting. She subsequently became the AB representative on the Medals and Award Committee.

Three special sessions were conducted at the 133rd meeting at Penn State University. Peter Tyack from Woods Hole Oceanographic Institution organized a session on ''Animal Communication.'' The session consisted of three invited papers. Daniel Raichel from Cooper Union organized a session on ''Bioacoustic Sensing of the Environment.'' Two invited papers and four contributed papers were presented at this session. The third session on ''Low-frequency Bioacoustics,'' was organized by Adam Frankel of Cornell Laboratory of Ornithology. Two invited presentations and six contributed papers were delivered at this session. Sam Ridgway from the Naval Command Control and Ocean Surveillance Center, RDT&E Division was elected a Fellow of the ASA at this meeting.

The special sessions for the 134th meeting in San Diego were finalized. There will be four special sessions. Sam Ridgway is organizing a named session: C. Scott Johnson session on whales and dolphin acoustics. Whitlow Au and Ashley Walker from University of Edinburgh are organizing a session on ''Biologically Inspired Acoustic Models and Systems.'' Peter Narins from UCLA and Ann Bowles are organizing a session on ''Acoustics of Reptiles and Amphibians,'' and Mardi Hastings of the Ohio State University and Charles Greene from Greeneridge Sciences Inc. are organizing a session on ''Instrumentation in Animal Bioacoustics.''

The AB committee obtained $500 from technical initiative funding for two student paper awards and will also join with P&P, Speech, and Music in conducting a student social at the San Diego meeting.

Four special sessions were also finalized for the joint ASA/ICA meeting in Seattle. There will be a session on ''Insect Acoustics,'' that will be organized by Timothy Forrest of Cornell University, a session on ''Avian Acoustics,'' that will be organized by Robert Benson of Texas A&M Uni-

versity, a session on ''Temporal Patterns and Rhythm in Bioacoustics,'' that will be organized by Christopher Clark of Cornell University, and a co-sponsored session with Acoustic Oceanography on ''Acoustics of Fish, Fisheries, and Plankton,'' that will be organized by Vance Holliday of Tracor Inc.

WHITLOW W. L. AU
Chair

Musical Acoustics

The 1996–97 year for the Technical Committee on Musical Acoustics (TCMU) has been an active period, encompassing a large number of special sessions at the ASA meetings, new directions in student involvement and in setting up a Web site, and much planning of future meetings.

The joint meeting of the Acoustical Society of Japan (ASJ) and the ASA in Honolulu in December 1996 offered a variety of special session topics. Ingo Titze and Seiji Niimi organized a special session on singing styles in the Pacific islands that included a topical focus on native Hawaiian music. Characterizing the international flavor of the meeting was a session organized by Isao Nakamura and Tom Rossing on traditional instruments of Japan and the Americas. This session was followed by a concert of multicultural traditional music that was enjoyed by all. A session on performance in music was organized by Kengo Ohgushi and Ed Carterette that combined both physical and subjective musical acoustics. Julius Smith and Tomoyasu Taguti chaired a lively session (co-sponsored by Signal Processing) on electronic synthesis and analysis, which featured research in such areas as physical modeling and instrument acoustics. Technical initiative funds were used to support an invited lecture by Kati Szego on Hawaiian vocal music.

The June 1997 meeting of the ASA in State College was an unusually large meeting for musical acoustics. Bob Pyle and Peter Hoekje organized a session on the lip reed and brass instruments that focused on scientific issues but considered also the art and craft of instrument-making. A special session on the acoustics of bells organized by Tom Rossing included two components supported by technical initiative funds: a presentation by Yang-Hann Kim on Korean bells and a handbell concert by the Westminster Concert Bell Choir directed by Kathleen Ebling-Thorne. In a pair of special sessions co-sponsored with the Committee on Education in Acoustics, Bob Collier and Ron Roy organized a broad view of the role of musical acoustics in teaching acoustics, mathematics, and engineering. David Butler chaired a special session on the convergence of music cognition and music theory, a topic at the intersection of the science and art of music. A memorial session in honor of W. Dixon Ward was sponsored by Psychological and Physiological Acoustics (P&P), Musical Acoustics, and Noise, and organized by Chris Turner. The session surveyed Dix's broad contributions to acoustics, and was appropriately concluded with the musical sounds of a barbershop quartet comprised of his acoustical and musical colleagues. We all benefit from the efforts contributed by these organizers of special sessions in Honolulu and State College.

New technical initiatives to increase student involvement in musical acoustics were begun at the State College meeting. A contest was organized for the best student presentation in musical acoustics. Among a number of excellent presentations, the winner chosen was Lucille Rossi, who presented research undertaken with Gerald Girolami entitled ''Use of note partials' energy to improve the identification of polyphonic piano signals'' [J. Acoust. Soc. Am. **101**, 3167(A) (1997)]. A second event was a very successful graduate student social that was co-sponsored with P&P, enabling students to informally meet ASA members and other students. An ongoing technical initiative that continued this past year supported a pair of educational workshops on acoustics for teachers conducted by Uwe Hansen.

This year was one of transition for the TCMU, as Uwe Hansen completed an effective and dynamic term as Chair at the beginning of the year. Our thanks go to Uwe and other members of the TCMU who have completed their terms—Donald Hall, Gabriel Weinreich, and Bob Collier as our representative on the Membership Committee. Musical acoustics has been enthusiastically international in scope and participation for some time, which is consistent with recent Technical Council initiatives to broaden international participation on Technical Committees. Thus, we are pleased to report that new incoming members joining the TCMU in the coming year include René Caussé (France), Adrianus Houtsma (The Netherlands), and Shigeru Yoshikawa (Japan). Completing the list of new members are James Beauchamp, W. Jay Dowling, Bruce Lawson, James Pyne, and William Hartmann, who also begins his term as Vice President-Elect of the ASA.

New *ex officio* members include Ian Lindevald (ASACOS) and Uwe Hansen (Membership). Thanks to Bill Hartmann, the TCMU now has a

home page on the WWW which can be accessed from the ASA home page. Among other things, the TCMU home page contains information on upcoming musical acoustics conferences, and an international list of musical acousticians with their areas of interest and/home-page addresses. Any musical acoustician may join this list by contacting the TCMU Chair.

The lists of special sessions for the San Diego meeting in Dec. 1997 and the joint meeting in Seattle in June 1998 of the International Congress on Acoustics (ICA) and the ASA have been finalized. The planning of special sessions for the joint meeting of the ASA with the European Acoustics Association in Berlin, Germany (Mar. 1999) must be completed by the San Diego meeting, and we will also make plans for the upcoming Oct. 1998 meeting in Norfolk.

Special sessions for San Diego include Computer Jazz Improvisation, Performance Anxiety, Piano. A best student presentation contest is scheduled for San Diego, as is a graduate and undergraduate student social cohosted with P&P and Animal Bioacoustics. Special sessions for Seattle include Timbre of Musical Sounds, Bowed String Instruments, Physics and Materials of Musical Instruments, Purposeful Use of Distortion in Musical Performance: The Eric Dolphy/Jimi Hendrix Celebratory Session, and Signal Modeling in Music Synthesis and Processing. The TCMU is co-sponsoring special sessions in Seattle on Opera Halls and Opera Singers (with Architectural Acoustics), and on Soundscapes—''Acoustical Landscapes'' in Natural and Built Environments (with Noise). Tentative special session topics for Norfolk include a session on Acoustics and Multimedia (with Signal Processing), and on the Organ and Performance Spaces (with Architectural Acoustics). Tentative plans for Berlin include Psychoacoustics and Audio Technology, and Musical Instruments and Structural Vibration (with Structural Acoustics and Vibration).

A major Technical Initiative this year is the co-sponsorship of the International Symposium on Musical Acoustics 1998 (ISMA '98) by the Catgut Acoustical Society and the ASA, with the theme ''Tone and Technology in Musical Acoustics.'' This satellite symposium on musical acoustics will immediately follow the ICA/ASA Seattle meeting from 26 June–1 July 1998 in Leavenworth, WA. ISMA '98 will bring together international leaders in the field and will provide instrument makers, musical acoustics researchers, and students the opportunity to interact in the more informal setting of a beautiful mountain retreat within driving distance of Seattle. The abstract deadline is 15 Dec. 1997; the Home Page of ISMA '98 at http://www.boystown.org/isma98/ provides up-to-date information.

DOUGLAS H. KEEFE
Chair

Psychological and Physiological Acoustics

This was an unusually busy year for Psychological and Physiological Acoustics (P&P), with two seccessful meetings, three changes in Associate Editors, planning for multiple future meetings, and a number of new initiatives. Attendance for recent meetings confirms that the Spring meeting is now the main meeting for P&P. Compared to meetings at these same locales roughly 10 years ago, attendance at Hawaii was notably lower whereas attendance at Penn State increased. A formal vote of the Committee reaffirmed the majority opinion to avoid special sessions for Fall meetings, with Hawaii always being an exception. Cosponsored with our Japanese colleagues, two excellent special sessions were held in Hawaii. One, organized by Hideki Kawahara and William Hartmann, was on auditory organization and representation. The second, organized by Toshio Sone and William Yost was on frequency analysis and timbre. Toshio Sone and William Yost organized the technical program. At the Spring meeting at Penn State, the technical program was organized by (P&P's) Richard Stern, and featured three special sessions. The first was our 1997 Technical Initiative for emphasizing a topic in physiological acoustics. The session presented physiological talks interwoven with related psychophysical talks on the processing of amplitude modulation. Neal Viemeister organized the session and helped fuel the lively discussion following the presentations. The second special session honored W. Dixon Ward, and was jointly sponsored with the Musical Acoustics and Noise Technical Committees. Our thanks to Christopher Turner, who organized the session, and to the former students, colleagues, and friends of Dix who presented a mixture of personal memories and science. As would have pleased Dix, the session closed with selections by a barbershop quartet comprised of Ira Hirsh, Mead Killion, Daniel Martin, and Thomas Rossing. The third special session explored a broad set of issues on auditory system plasticity and perceptual learning. These ranged from anatomical and physiological evidence of plasticity to human perceptual learning with cochlear implants and infant language acquisition. The

session was well attended despite being the final P&P session for the week. An unexpected highlight of the week was a graduate-student reception, hosted jointly by P&P and Musical Acoustics. P&P has not tried such a reception for some years, but both the students and members attending felt it was successful and should definitely be continued at future meetings. Results of the ASA elections were announced, and P&P congratulates William Hartmann (VP-Elect) and Janet Weisenberger (Executive Council).

A more general membership initiative this year was a letter campaign to nudge our associate members to upgrade their status to full membership. About 140 letters were sent out and 30 upgrades were received. Several other new initiatives focused on better email and Internet communication with our members. WWW pages for P&P linked to ASA's home page are online, thanks to Stephen Neely. Check them out and contact either Steve (neely@boystown.org) or Brent Edwards (bedwards@compuserve.com) to place new information there. Electronic submission of abstracts for the meetings was widely used and members strongly supported the process. A large email list was developed, with the help of Elaine Moran, of ASA members listing P&P as their primary affiliation. Our first attempts to use this list to encourage voting in the ASA elections and attendance at the Penn State meeting worked quite well. We expect to continue to update and develop these lists to pass along news and poll the membership on issues that arise. Instead of starting our own listserver, the members attending the business meeting at Penn State voted to encourage members to join another active server in our area, the Auditory List.

Publication in JASA remained strong for P&P; with the number of papers in physiological acoustics increasing beyond a reasonable workload for one associate editor. Consequently, we were pleased when Daniel Martin, Editor-in-Chief, agreed to add Brenda Lonsbury-Martin as a second associate editor in physiological acoustics, beginning in January 1997. Two of the three associate editors in psychological acoustics will be ending their three-year terms this summer. We thank Raymond Dye and Walt Jesteadt for their hard work, and welcome Wesley Grantham and Robert Shannon as their replacements. At the business meeting, there was much discussion of the proposed electronic publishing options for circulating preprints and publishing short articles. Despite various concerns raised about these experimental ventures, the members remain strongly supportive of moving forward with electronic publishing, particularly handling as much of the current review process as possible electronically.

Like all the other Technical Committees, much time this year was spent planning special sessions for early deadlines for future meetings. Primarily because of the Fall-meeting factor, the membership decided not to hold a special session on hearing impairment to complement a workshop to be offered by Speech this Fall at San Diego. However, any papers received by P&P on the topic will be organized for the day following the speech workshop to encourage our interested members to attend. For the Seattle 1998 meeting, we decided on two special sessions. One will be on auditory attention, to be organized by Ervin Hafter, and the second on auditory development, to be organized by Lynne Werner. Six tentative sessions were proposed for the Berlin 1999 meeting on topics likely to be of mutual interest to us and our European colleagues: cross-spectral auditory processing, recent advances in models of auditory processing, peripheral compression in hearing for normal and impaired ears, localization and speech perception in realistic complex acoustic environments, and otoacoustic emisisons. In addition, there are tentative plans for a 1-day workshop at Berlin to be held jointly with a non-ASA Auditory Display group. Members interested in these or other potential topics should contact me. The annual election was held at Penn State for new members of the Technical Committee for 1998. The outgoing members, chaired by Glenis Long and Evan Relkin, prepared the ballot and the six elected were Prudence Allen, Robert Carlyon, Brent Edwards, Thomas Hanna, Nina Kraus, and John Rosowski. The members elected last spring who now go on the committee are David Eddins, Mary Florentine, John Grose, Marjorie Leek, Elizabeth Strickland, and Fan-Gang Zeng. Finally, thanks to the outgoing members for their work these past three years (Glenis Long, Evan Relkin, Brian Moore, John Middlebrooks, Bradford May, and Gregory Wakefield), and to the many P&P members who served on committees, chaired sessions and generally helped get things done.

DONNA L. NEFF
Chair

Signal Processing in Acoustics

The Interdisciplinary Technical Group on Signal Processing in Acoustics (SP) has seen continued growth in its first few years. It has held its fifth and sixth meetings at the Honolulu and Penn State meetings of the ASA.

At the Honolulu meeting SP was lead sponsor on five special session and three regular sessions, as well as co-sponsor on three special sessions. A total of 36 invited papers and 99 contributed papers were presented, with co-sponsorship from AA, AO, EA, PA, SC, and MU Technical Committees. Yoshio Yamasaki and David Havelock represented SP at the TPOM for Honolulu. The ITG-SP met on Tuesday evening, 3 December (day 2 of the Honolulu meeting) at 7:00 p.m. with about 20 people attending. Voting procedures for election of the Chair of SP were established, with eligibility to vote based on attendance at an ITG meeting in the past 3 years, excluding the meeting at which the ballots are counted.

The Gallery of Acoustics, with Randall W. Smith as curator, was a popular addition to the Honolulu meeting. It consisted of four still images and four videos, as well as an encore section of past winners. The winning entries at the Honolulu Gallery of Acoustics were: ''Structural and Acoustic Response to Stiffened Plates'' (First Place and Peoples' Choice Award winner), a still entry submitted by Benjamin Cray; ''Acoustics and Vibration Characteristics of the King Seong-Deok Bell'' (Second Place winner), a video submitted by Yang-Hann Kim; ''Plate Mode Visualization with Laser-modulated Phase-stepped Digital Shearography'' (Third Place winner), a still entry submitted by Benjamin A. Bard, Shudong Wu, and Giro-wen Lu. The remaining entries to the Gallery of Acoustics were: ''Free Oscillations and Surfactant Studies of Superdeformed Drops in Microgravity,'' a still entry by R. Apfel, Y. Tian, J. Jankovsky, X. Chen, T. Shi, R. G. Holt, E. Trinh, A. Croonquist, K. Thornton, A. Sacco, Jr., C. Coleman, F. Leslie, and D. Matthiesen; ''Ultrasonic Wave Visualization with Laser-modulated Phase-stepped Digital Shearography,'' a still entry by Benjamin A. Bard, Grant A. Gordon, and Shudong Wu; ''Tomographic Acoustic Microscopy of a Penny'' (late arrival), a still entry by S. Davis and Hua Lee; ''Synthetic-aperture Sonar Image of a Sunken Airplane'' (late arrival), a still entry by Brent L. Douglas and Hua Lee; ''Microphones for the Multitudes,'' a still entry by David Havelock; ''Sound Field Fluctuations due to Atmospheric Turbulence,'' a video entry by David Havelock; ''Observation of Micro-bubble Oscillations,'' a video entry by J. Jankovsky, Y. Tian, J. Ketterling, and R. Apfel.

At the Penn State meeting, SP was lead sponsor on three special sessions and three regular sessions, as well as co-sponsor on four special or regular sessions. A total of 22 invited papers and 77 contributed papers were presented, including one distinguished lecture. Sessions were co-sponsored with EA, NS, PA, SA, SC, UW, and Noise-Con. David Havelock represented SP at the TPOM for Penn State. The ITG-SP met on Wednesday evening, 18 June (day 3 of the Penn State meeting) at 7:00 p.m. with about 20 people attending. An election was held, with James Candy and Gary Wilson running as candidates for Chair of SP. Twenty four ballots were received, at the meeting or by prior mail, from attendees of previous ITG-SP meeting. A draw vote was resolved when James Candy stepped down in favor of Gary Wilson, who will take office as the new Chair of SP at the end of the San Diego meeting (1–5 December 1997). Executive Council has approved the continuance of the ITG-SP for another three years, on the unanimous recommendation of the Technical Council.

The new nominees to the committee of the Interdisciplinary Technical Group on Signal Processing in Acoustics, with terms to 2000, are John C. Burgess, Christian P. deMoustier, Gary W. Elko, Stergios Stergiopoulos, Julius O. Smith III, and Yoshio Yamasaki.

The ITG-SP has had technical initiatives for the Gallery of Acoustics, co-sponsorship of the 23rd International Symposium on Acoustic Imaging in Boston, demonstrations in computational acoustics at Honolulu, and an invited speaker at Penn State.

Future activities of the ITG-SP will see a continued effort to present signal processing in acoustics in an interdisciplinary context, as in the future editions of the Gallery of Acoustics and a planned series of special sessions on Acoustics in Multimedia.

DAVID I. HAVELOCK
Chair

Speech Communication

The Technical Committee on Speech Communication sponsored one workshop and two special sessions at the Indianapolis Meeting in May. The workshop, entitled ''Informal Workshop on Voice Perception,'' was organized by David Pisoni. It was the first such workshop sponsored by SC at an ASA meeting. It was well attended and generally perceived as a very successful enterprise. The special sessions were ''Speech and Music: Exchange of Ideas Methods and Findings'' (sponsored jointly with Musical Acoustics) and ''Pediatric Cochlear Implants: Recent Studies of Speech Perception and Production.'' Both sessions were well attended.

We sponsored (all jointly with our ASJ counterparts) five special sessions at the Honolulu Joint meeting. These covered a wide range of topics and were both well attended and well received. Titles of the sessions were: ''Speaker Characteristics: Effects of Gestural Variability on Speech Perception,'' ''Speech Communication for the Next Decade: New Directions of Research Technological Development and Evolving Applications,'' ''First and Second Language Acquisition: Facts and Models,'' ''Robust Speech Recognition,'' and ''Prosody Modeling and its Application to Recognition and Synthesis.''

As has been customary in the last several years, the Speech TC sponsored graduate student receptions and student paper competitions at both meetings. These programs are well established and generally viewed as highly successful and worth continuing by the committee.

TERRANCE M. NEAREY
Chair

Members of Technical and Administrative Committees and Technical Groups of the Acoustical Society of America

The Technical and Administrative Committees and Technical Groups listed below have been appointed by the Executive Council. These appointments, with such changes as may be made by the President from time to time, will be in effect until the Spring meeting of the Society in 1998.

Technical Committees 1997–1998

Acoustical Oceanography

	Term to
Darrell R. Jackson, *Chair* to 1998	1998
Robert W. Farwell	2000
Gary J. Heald	2000
Anatoliy N. Ivakin	2000
Subramanian D. Rajan	2000
Barbara J. Sotorin	2000
Ralph A. Stephen	2000
Alexander G. Voronovich	2000
Suk Wang Yoon	2000
Michael G. Brown	1999
Christian P. de Moustier	1999
James A. Mercer	1999
Marshall H. Orr	1999
Er-Chang Shang	1999
Robert D. Stoll	1999
Dajun Tang	1999
Alexandra I. Tolstoy	
Mohsen Badiey	1998
Nicholas P. Chotiros	1998
James F. Lynch	1998
Hasan Oguz	1998
John R. Proni	1998
Michael D. Richardson	1998
Tokuo Yamamoto	1998

Ex officio:

Jeffrey H. Krolik, Associate Editor of JASA
David L. Bradley, Associate Editor of JASA
Stanley L. Chin-Bing, Associate Editor of JASA
D. Vance Holliday, member of Medals and Awards Committee
Michael J. Buckingham, member of Membership Committee
Alexandra I. Tolstoy, member of ASACOS

1260   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Acoustical News—USA   1260

Engineering Acoustics

	Term to
Thomas R. Howarth, *Chair* to 2000	2000
Kim C. Benjamin	2000
Allan C. Cummings	2000
Gary W. Elko	2000
Robert D. Finch	2000
Guillermo C. Gaunaurd	2000
Gordon Hayward	2000
Dehua Huang	2000
Sung Hwan Ko	2000
Victor Nedzelnitsky	2000
Ahmet Selamet	2000
James E. West	2000
Oscar B. Wilson	2000
George S. K. Wong	2000
Steven R. Baker	1999
David A. Brown	1999
Stephen C. Butler	1999
W. Jack Hughes	1999
K. Kirk Shung	1999
R. Lowell Smith	1999
Stephen C. Thompson	1999
Arnie L. Van Buren	1999
A. Mark Young	1999
Mahlon D. Burkhard	1998
Fred G. Geil	1998
Robert Hickling	1998
Dennis F. Jones	1998
Jan F. Lindberg	1998
Alan Powell	1998
Roger T. Richards	1998
Harold C. Robinson	1998
Kenneth D. Rolt	1998
Ludwig W. Sepmeyer	1998
Allan J. Zuckerwar	1998

Ex officio:
Henry E. Bass, Associate Editor of JASA
Stanley L. Ehrlich, Associate Editor of JASA
Robert D. Finch, member of Medals and Awards Committee
Sung H. Ko, member of Membership Committee
Mahlon D. Burkhard, member of ASACOS

Musical Acoustics

	Term to
Douglas H. Keefe, *Chair* to 1999	1999
James W. Beauchamp	2000
Rene E. Causse	2000
W. Jay Dowling	2000
William M. Hartmann	2000
Adrianus J. Houtsma	2000
Bruce A. Lawson	2000
James M. Pyne	2000
Shigeru Yoshikawa	2000
George A. Bissinger	1999
James P. Cottingham	1999
Uwe J. Hansen	1999
Peter L. Hoekje	1999
Ian M. Lindevald	1999
Daniel W. Martin	1999
Judith C. Brown	1998
Carleen M. Hutchins	1998
Max V. Mathews	1998

Thomas D. Rossing	1998
Julius O. Smith	1998
William J. Strong	1998

Ex officio:
Stanley L. Ehrlich, Associate Editor of JASA
William J. Strong, Associate Editor of JASA
Thomas D. Rossing, member of Medals and Awards Committee
Uwe J. Hansen, member of Membership Committee
Ian M. Lindevald, member of ASACOS

Noise

	Term to
Richard J. Peppin, *Chair* to 2000	2000
Keith Attenborough	2000
Sergio Beristain	2000
Robert D. Bruce	2000
John C. Burgess	2000
Angelo J. Campanella	2000
Robert J. Comparin	2000
T. James DuBois	2000
John J. Earshen	2000
Tony F. W. Embleton	2000
John Erdreich	2000
B. John Feng	2000
Robert D. Hellweg	2000
Robert M. Hoover	2000
Daniel L. Johnson	2000
Tor S. D. Kihlman	2000
Sonoko Kuwano	2000
Chantal Laroche	2000
Robert Lotz	2000
Lisa Louie	2000
Alan H. Marsh	2000
Ralph T. Muehleisen	2000
Kenneth J. Plotkin	2000
Joseph Pope	2000
Julia D. Royster	2000
Brigitte Schulte-Fortkamp	2000
Jean Tourret	2000
Robert W. Young	2000
Elliott H. Berger	1999
Bennett M. Brooks	1999
Lawrence S. Finegold	1999
Jerry G. Lilly	1999
George A. Luz	1999
Luc Mongeau	1999
Mary M. Prince	1999
Jack E. Randorff	1999
Nancy S. Timmerman	1999
Laura A. Wilber	1999
Martin Alexander	1998
Arno S. Bommer	1998
Paul R. Donavan	1998
Larry H. Royster	1998
John P. Seiler	1998
Noral D. Stewart	1998
Michael R. Stinson	1998
Alice H. Suter	1998
Louis C. Sutherland	1998
Jiri Tichy	1998

Ex officio:
Michael R. Stinson, Associate Editor of JASA
Stanley L. Ehrlich, Associate Editor of JASA
Elliott H. Berger, member of Medals and Awards Committee
Daniel L. Johnson, member of Membership Committee
Arno S. Bommer, member of ASACOS

1262   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Acoustical News—USA   1262

Structural Acoustics and Vibration

	Term to
Jerry H. Ginsberg, *Chair* to 2000	2000
Alain C. Berry	2000
Hunter C. Cohen	2000
David Feit	2000
Guillermo C. Gaunaurd	2000
Karl Grosh	2000
Jean L. Guyader	2000
Sabih I. Hayek	2000
Francis Kirschner	2000
Mauro Pierucci	2000
Andrew F. Seybert	2000
Scott D. Sommerfeldt	2000
Vasundara V. Varadan	2000
Richard D. Vogelsong	2000
John A. Wilder	2000
John A. Burkhardt	1999
Courtney B. Burroughs	1999
Joseph M. Cuschieri	1999
Alison Flatau	1999
Richard F. Keltie	1999
J. Adin Mann III	1999
Jean R. Nicolas	1999
Allan D. Pierce	1999
Victor W. Sparrow	1999
J. Stuart Bolton	1998
Christopher R. Fuller	1998
Joel M. Garrelick	1998
Louis A. Herstein III	1998
Geoffrey L. Main	1998
John J. McCoy	1998
Aynur Unal	1998
Eric E. Ungar	1998
Vijay K. Varadan	1998
Earl G. Williams	1998

Ex officio:
Courtney B. Burroughs, Associate Editor of JASA
Paul J. Remington, Associate Editor of JASA
Allan D. Pierce, member of Medals and Awards Committee
Courtney B. Burroughs, member of Membership Committee
Louis A. Herstein, member of ASACOS

Underwater Acoustics

	Term to
George V. Frisk, *Chair* to 2000	2000
David L. Bradley	2000
Curtis I. Caldwell	2000
William M. Carey	2000
Nicholas P. Chotiros	2000
Donald R. Del Balzo	2000
Frederick R. DiNapoli	2000
David R. Dowling	2000
Robert W. Farwell	2000
Roger C. Gauss	2000
Darrell R. Jackson	2000
Finn B. Jensen	2000
Roger W. Meredith	2000
John R. Preston	2000
Henrik Schmidt	2000
Ralph A. Stephen	2000
Robert D. Stoll	2000
Frederick D. Tappert	2000
David J. Thomson	2000
Alexander G. Voronovich	2000

R. Jeffrey Wilkes	2000
Tsih C. Yang	2000
Robert A. Zingarelli	2000
Ralph N. Baer	1999
Michael K. Broadhead	1999
Shira L. Broschat	1999
Berlie A. Brunson	1999
Peter G. Cable	1999
Marshall V. Hall	1999
Samuel W. Marshall	1999
Marshall H. Orr	1999
Dan J. Ramsdale	1999
Timothy H. Ruppel	1999
Natalia A. Sidorovskaia	1999
Kevin B. Smith	1999
Dajun Tang	1999
Christopher T. Tindle	1999
Alexandra I. Tolstoy	1999
Homer P. Bucker	1998
Stanley A. Chin-Bing	1998
N. Ross Chapman	1998
Michael D. Collins	1998
Christopher Feuillade	1998
Ralph R. Goodman	1998
Diana F. McCammon	1998
Stephen K. Mitchell	1998
Guy V. Norton	1998
Peter M. Ogden	1998
Eric I. Thorsos	1998
Michael F. Werby	1998

Ex officio:
Jeffrey L. Krolik, Associate Editor of JASA
David L. Bradley, Associate Editor of JASA
Stanley A. Chin-Bing, Associate Editor of JASA
Ralph R. Goodman, member of Medals and Awards Committee
Peter H. Rogers, member of Membership Committee
Arnie L. Van Buren, member of ASACOS

Signal Processing in Acoustics
Interdisciplinary Technical Group

	Term to
David I. Havelock, *Chair*	2000
John C. Burgess	2000
Christian P. deMoustier	2000
Gary W. Elko	2000
Stergios Stergiopoulos	2000
Julius O. Smith III	2000
Yoshio Yamasaki	2000
Edith L. R. Corliss	1999
John M. Noble	1999
James C. Rogers	1999
David C. Swanson	1999
P. G. Vaidya	1999
Gary R. Wilson	1999
J. Stuart Bolton	1998
Stanley L. Ehrlich	1998
Richard M. Guernsey	1998
Joseph Pope	1998
James G. Ryan	1998
Randall W. Smith	1998
Scott A. Van Duyne	1998
William A. Von Winkle	1998

Ex officio:
Jeffrey L. Krolik, Associate Editor of JASA

Administrative Committees 1997–1998

Archives and History

	Term to
Edith L. R. Corliss, *Chair* to 2000	2000
Richard K. Cook	2000
John W. Kopec	2000
Richard J. Peppin	2000
William J. Cavanaugh	1999
Carleen M. Hutchins	1999
Henry E. Bass	1998
Malcolm J. Crocker	1998

Books[+]

	Term to
Mohsen Badiey, *Chair* to 1999	1999
Robert D. Frisina	2000
Jerry H. Ginsberg	2000
Philip L. Marston	2000
Joseph Pope	2000
Robert A. Walkling	2000
Stephen N. Wolf	2000
Suzanne Wolf	2000
Nancy S. McGarr	1999
Jeffrey A. Nystuen	1999
Neil A. Shaw	1999
Emily A. Tobey	1999
Victor W. Sparrow	1998

Ex officio:
James F. Bartram, Associate Editor of JASA for Book Reviews

Development

	Term to
Mahlon D. Burkhard, *Chair* to 1998	1998
D. Kent Lewis	2000
John V. Bouyoucos	1999
Francis Kirschner	1998

Ex officio:
William W. Lang, Treasurer, *ex officio*
Charles E. Schmid, Executive Director, *ex officio*

Education in Acoustics

	Term to
Victor W. Sparrow, *Chair* to 2000	2000
Anthony A. Atchley	2000
Fredericka Bell-Berti	2000
E. Carr Everbach	2000
Uwe J. Hansen	2000
Elizabeth S. Ivey	2000
P. K. Raju	2000
Daniel A. Russell	2000
Emily A. Tobey	2000
D. Michael Daly	1999
Logan E. Hargrove	1999
Mardi C. Hastings	1999
Michel T. T. Jackson	1999
Murray F. Korman	1999
Diana F. McCammon	1999
James M. Sabatier	1999
Joseph R. Sears	1999

James E. West	1999
Wayne M. Wright	1999
Michael K. Wynne	1999
Steven M. Haas	1998
Philip L. Marston	1998
Sally G. Revoile	1998
Thomas D. Rossing	1998
Ronald A. Roy	1998
Linda K. Thibodeau	1998

Investments

	Term to
John V. Bouyoucos, *Chair* to 1998	1998
Leo L. Beranek	1998
Kenneth M. Eldred	1998

William W. Lang, Treasurer, *ex officio*

Long Range Planning

	Term to
Anthony A. Atchley, *Chair* to 2000	2000
Dana S. Hougland	2000
Scott D. Sommerfeldt	2000
Murray Strasberg	2000
Floyd Dunn	1999
Mardi C. Hastings	1999
Stephen C. Thompson	1999
Randall W. Smith	1998
Sigfrid D. Soli	1998
Jiri Tichy	1998

James E. West, President-Elect, *ex officio*

Medals and Awards

		Term to
Sabih I. Hayek, *Chair* to 1998		1998
Anthony A. Atchley	Physical Acoustics	2000
Elliott H. Berger	Noise	2000
Mardi C. Hastings	Animal Bioacoustics	2000
Wesley L. Nyborg	Bioresponse to Vibration/ Biomedical Ultrasound	2000
William A. Yost	Psychological and Physiological Acoustics	2000
Steven M. Brown	Architectural Acoustics	1999
Ralph R. Goodman	Underwater Acoustics	1999
Allan D. Pierce	Structural Acoustics and Vibration	1999
Robert D. Finch	Engineering Acoustics	1998
D. Vance Holliday	Acoustical Oceanography	1998
Thomas D. Rossing	Musical Acoustics	1998
Sigfrid D. Soli	Speech Communication	1998

Meetings June 1997–Dec. 1997

George C. Maling, *Chair* to 1999
Ilene J. Busch-Vishniac, Vice President
William M. Hartmann, Vice President-Elect
Uwe J. Hansen, Chair, Spring 1996
John C. Burgess, Chair, Fall 1996
Sabih I. Hayek, Chair, Spring 1997
Mauro Pierucci, Chair, Fall 1997
Lawrence A. Crum, Chair, Spring 1998
Kevin P. Shepherd, Chair, Fall 1998

Elaine Moran, ASA Office Manager, *ex officio*
Charles E. Schmid, Executive Director, *ex officio*

Subcommittee on Exhibits

Richard J. Peppin, *Chair*
Tony F. W. Embleton
Robert Finnegan
Noland D. Lewis
George C. Maling
Julia D. Royster

Membership

		Term to
Joseph W. Dickey, *Chair* to 2000		2000
Richard R. Fay	Psychological and Physiological Acoustics	2000
Uwe J. Hansen	Musical Acoustics	2000
Sung H. Ko	Engineering Acoustics	2000
Maureen Stone	Speech Communication	2000
Vacant	Animal Bioacoustics	2000
Anthony J. Brammer	Bioresponse to Vibration/ Biomedical Ultrasound	1999
Courtney B. Burroughs	Structural Acoustics and Vibration	1999
Burton G. Hurdle	Foreign Members	1999
Daniel L. Johnson	Noise	1999
Michael J. Buckingham	Acoustical Oceanography	1998
Richard M. Guernsey	Architectural Acoustics	1998
Peter H. Rogers	Underwater Acoustics	1998
Steven L. Garrett	Physical Acoustics	1998

Public Relations

	Term to
Paul A. Baxley, *Chair* to 2000	2000
Anthony J. Brammer	2000
William J. Cavanaugh	2000
T. James DuBois	2000
Blas Espinoza-Varas	2000
Holly S. Haggerty	2000
Helen Ann McCaffrey	2000
Victor Nedzelnitsky	2000
Mauro Pierucci	2000
Thomas D. Rossing	2000
Bor-Tsuen Wang	2000
Ewart A. Wetherill	2000
John Erdreich	1999
Lawrence L. Feth	1999
Logan E. Hargrove	1999
Burton G. Hurdle	1999
Carlos R. Jimenez-Dianderas	1999
James F. Lynch	1999
Joanne L. Miller	1999
Joseph Pope	1999
Punita G. Singh	1999
Barbara J. Sotirin	1999
Jamal Assaad	1998
E. Carr Everbach	1998
Christy K. Holland	1998
D. Kent Lewis	1998
David Lubman	1998
Deborah A. Summa	1998

Daniel W. Martin, Editor-in-Chief, *ex officio*
Elaine Moran, ASA Office Manager, *ex officio*
Charles E. Schmid, Executive Director, *ex officio*
Thomas D. Rossing, Echoes Editor, *ex officio*

Publication Policy

	Term to
Floyd Dunn, *Chair* to 2000	2000
A. Simmons	2000
Robert C. Bilger	1999
James F. Lynch	1999
James H. Miller	1999
George S. K. Wong	1999
Allan J. Zuckerwar	1999
Walter G. Mayer	1998
Sigfrid D. Soli	1998
Michael R. Stinson	1998

Lawrence A. Crum, President-Elect, *ex officio*
Daniel W. Martin, Editor-in-Chief, *ex officio*

Regional Chapters

Roger T. Richards, *Chair* to 2000

Elmer L. Hixson	Austin
Angelo J. Campanella	Central Ohio
Courtney B. Burroughs	Central Pennsylvania
John W. Kopec	Chicago
Vacant	Cincinnati
Edwin H. Toothman	Delaware Valley
Gary W. Siebein	Florida
Timothy J. Foulkes	Greater Boston
Robert M. Hoover	Houston
Michael J. Anderson	Inland Northwest
Neil A. Shaw	Los Angeles
Hari S. Paul	Madras, India
Marehalli G. Prasad	Metropolitan New York
Elizabeth A. McLaughlin	Narragansett
Larry H. Royster	North Carolina
Vacant	Northern California
Paul D. Joppa	Northwest
Donald A. Murphy	Orange County
Mauro Pierucci	San Diego
David Braslau	Upper Midwest
Timothy S. Margulies	Washington, D.C.
Thomas M. Disch	Wisconsin

Victor W. Sparrow, Chair, Education in Acoustics, *ex officio*

Rules and Governance

	Term to
Tony F. W. Embleton, *Chair* to 2000	2000
Charles E. Schmid	2000
Elaine Moran	1999
Murray Strasberg	1998

Special Fellowships

	Term to
Wayne M. Wright, *Chair* to 1999	1999
James D. Miller	2000
James E. West	2000
Henrik Schmidt	1999
Constantine Trahiotis	1998
Joseph W. Dickey	1998

Standards

Executive Committee
Daniel L. Johnson, Chair (Standards Director)
Paul D. Schomer, Vice Chair
Avril Brenig, Standards Manager, *ex officio*

S1 Representation

John P. Seiler, Chair S1 and ASA representative on S1

George S. K. Wong, Vice Chair S1 and ASA alternate representative on S1

S2 Representation

David J. Evans, Chair S2 and ASA representative on S2

Richard F. Taddeo, Vice Chair S2 and ASA alternate representative on S2

S3 Representation

Lawrence S. Finegold, Chair S3 and ASA alternate representative on S3

R. F. Burkard, Vice Chair S3 and ASA representative on S3

S12 Representation

Paul D. Schomer, Chair S2 and ASA representative on S12

Bennett M. Brooks, Vice Chair S12

William J. Galloway, ASA alternate representative on S12

International TAGs (ex officio)

Paul D. Schomer, Chair, U.S. TAG for ISO/TC 43

Elliott H. Berger, U.S. TAG for ISO/TC 94/SC 12

David J. Evans, Chair, U.S. TAG for ISO/TC 108

Victor A. Nedzelnitsky, U.S. Technical Advisor for IEC/TC 29

ASA Technical Committee Representatives

Ilene J. Busch-Vishniac, Chair of ASA Technical Council, *ex officio*

Alexandra I. Tolstoy, Acoustical Oceanography

Ann E. Bowles, Animal Bioacoustics

George E. Winzer, Architectural Acoustics

Stanley J. Bolanowski, Jr., Bioresponse to Vibration/Biomedical Ultrasound

Mahlon D. Burkhard, Engineering Acoustics

Ian M. Lindevald, Musical Acoustics

Arno S. Bommer, Noise

Sameer I. Madanshetty, Physical Acoustics

Christine M. Rankovic, Psychological and Physiological Acoustics

Astrid Schmidt-Nielsen, Speech Communication

Louis A. Herstein, Structural Acoustics and Vibration

Arnie L. Van Buren, Underwater Acoustics

ASA Officers

William W. Lang, Treasurer, *ex officio*

Charles E. Schmid, Executive Director, *ex officio*

Past Chair of ASACOS (ex officio)

Tony F. W. Embleton

Associate Editors for Standards News—JASA (ex officio)

Avril Brenig

George S. K. Wong

ANSI Acoustical Standards Board (ex officio)

Avril Brenig, ASA Representative

Henning E. von Gierke, ASA Alternate Representative

Tutorials

	Term to
Joseph Pope, *Chair* to 2000	2000
Kenneth E. Gilbert	2000
Kenneth J. Plotkin	2000
Yves H. Berthelot	1999
Uwe J. Hansen	1999
Alexandra I. Tolstoy	1999
Charles E. Schmid, Executive Director, *ex officio*	

Women in Acoustics

	Term to
Barbara J. Sotirin, *Chair* to 1999	1999
Arlene E. Carney	2000
Chi-Fang Chen	2000
Peter H. Dahl	2000
Helen M. Hanson	2000
Elizabeth S. Ivey	2000
Bozena Kostek	2000

Dana S. Hougland	1999
Juliette Ioup	1999
Ellen S. Livingston	1999
Dianna F. McCammon	1999
Sally G. Revoile	1999
Victor W. Sparrow	1999
Ann E. Bowles	1998
Rhona P. Hellman	1998
Alexandra I. Tolstoy	1998
Janet M. Weisenberger	1998
James H. Miller	1998

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

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].
19–20 Sept.	Fifth Annual Conference on Management of the Tinnitus Patient, Iowa City, IA [Richard Tyler, Univ. of Iowa, Dept. of Otolaryngology–Head & Neck Surgery, 200 Hawkins Dr., C21-3GH, Iowa City, IA 52242-1078, Tel.: 319-356-2471; FAX: 319-353-7639; E-mail: rich-tyler@uiowa.edu].
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: hearingaid@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

9–13 Feb.	1998 Ocean Sciences Meeting, San Diego, CA [American Geophysical Union, 2000 Florida Ave., N.W., Washington, DC 20009, Tel.: 202-462-6900; FAX: 202-328-0566; WWW: http://www.agu.org].
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].
7–12 July	Vienna and the Clarinet, Ohio State Univ., Columbus, OH [Keith Koons, Music Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354, Tel.: 407-823-5116; E-mail: kkons@pegasus.cc.ucf.edu].
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].

1267   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Acoustical News—USA   1267

OBITUARIES

This section of the Journal publishes obituaries concerning the death of Fellows of the Society and other acousticians eminent in the world of acoustics. When notified, the Editor-in-Chief solicits a summary of the person's life and contributions from an ASA member thoroughly familiar with the details, if possible. If a promised obituary is never received, a brief obituary notice may be published later.

Robert J. Urick • 1915–1996

Robert J. Urick, a Fellow of the Society and the 1988 recipient of the Society's ''Pioneers of Underwater Acoustics'' Medal, died 31 December 1996 of multiple myeloma. Born in Brooklyn, New York, Bob received his BS degree from Brooklyn College in 1935. He performed subsequent graduate work at Columbia University and the California Institute of Technology, where he received his MS in 1939.

Robert Urick started his career at Shell Oil and Texas Company working on seismic exploration from 1937 through 1942. It was a natural transition to underwater sound when, in 1942, he joined the staff of U.S. Navy Radio and Sound Laboratory in San Diego. In 1945 Bob transferred to the Naval Research Laboratory in Washington, DC. From 1955 through 1957 Bob was a Professor of Engineering Research at the Penn State Ordnance Research Laboratory. In 1957 he became the head of Research Division at the Navy Mine Defense Laboratory in Panama City, Florida. In 1960 he returned to Washington to the Naval Ordnance Laboratory in Silver Spring, Maryland from which he retired in 1975. Following retirement he was an adjunct professor at Catholic University of America.

During his career Bob received numerous other awards, including the Distinguished Technical Achievement Award of Oceanic Engineering Society and the Distinguished Civilian Service Award from the U.S. Navy. Bob also served as an associate editor of this *Journal* from 1975 to 1978.

As a government scientist, his accomplishments span the entire spectrum of underwater acoustics. His work for the U.S. Navy provided the foundation for countless designs, modifications, and improvements. He was the author of the well-known book *Principles of Underwater Sound* and two others, *Propagation of Sound in the Sea* and *Ambient Noise in the Sea*, as well as over 200 papers and technical reports. Bob also taught traditional graduate school courses and short courses in underwater sound.

On a personal note, it was through one of Bob's short courses that one of us (Miller), as a young engineer, decided to pursue a career in underwater acoustics. One of us (Gaunaurd), a co-worker and co-author with Bob, took a year-long graduate course in ''Underwater Sound'' that Bob taught at Catholic University in the sixties, and Bob was his first supervisor at the Naval Ordnance Laboratory.

Throughout his lengthy career Bob had a great impact not only on naval systems, but also on the younger group of people that either studied or worked with him at the various institutions with which he was associated. His passing is a great loss for the acoustical scientific community.

Bob's wife of 42 years, Julia, died in February of 1996. He is survived by three children: Maureen J. Bledsoe of Silver Spring, Victoria Cole of San Diego, Robert M. Urick of Crownsville, Maryland and seven grandchildren.

GUILLERMO C. GAUNAURD
Naval Surface Warfare Center, Carderock Division
Silver Spring, Maryland

JAMES H. MILLER
University of Rhode Island
Narragansett, Rhode Island

Hans L. Oestreicher • 1912–1996

Hans Oestreicher, a Fellow of the Acoustical Society, died unexpectedly due to heart failure on 18 July 1996. For the last dozen years, he had been enjoying his retirement, his family, and his travels all over the globe. Acoustics and the Acoustical Society lost a productive researcher and thinker.

Hans received the Ph.D. degree in mathematics and an advanced degree in physics at the University of Vienna, Austria, in 1934. Later he was a research assistant there, and taught graduate courses in mathematics until he joined the Helmholtz Institute in Bavaria as a research mathematician, later becoming Chief of the Applied Mathematics Branch. He came to the United States in 1947. After a few years he headed the Mathematics and Analysis Branch of the Bioacoustics Division of the Aerospace Medical Research Laboratory at Wright–Patterson Air Force Base, Dayton, Ohio. For 35 years he conducted and directed research in theoretical acoustics and biophysics. He made lasting contributions to our theoretical knowledge in three distinctly different areas. The first was the sound radiation from moving sound sources. His original theoretical predictions were way ahead of the experiments, which later confirmed them. The second area concerned the propagation of sound and vibration in a viscoelastic medium, such as human tissue. This classical work provided the foundation and explanation for many observations in physiology and medicine and contributed to the emerging biological use of ultrasound. The third area, in which he became more and more interested in later years, concerned the mathematical understanding and description of the capabilities of the human brain, its reasoning, decision making, and intelligence: human intelligence compared to machine intelligence. The word Bionics, today used and misused by the whole world, was first coined in Hans' group to describe the study of life-like systems.

Hans was also a Senior Member of the IEEE, a member of the New York Academy of Sciences, and of the American Mathematical Society. He was associate editor of ''IEEE Transactions on Systems, Man and Cybernetics,'' and of the Journal of ''Mathematical Modeling.'' Hans Oestreicher had a full, productive and happy life. His knowledge, curiosity, and humor were enjoyed thankfully by all who worked with him or were guided by him on their way to becoming theoretical acousticians.

HENNING E. VON GIERKE

BOOK REVIEWS

James F. Bartram

94 Kane Avenue, Middletown, Rhode Island 02842

The opinions expressed are those of the individual reviewers and are not necessarily endorsed by the Editorial Board of this Journal.

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

Chaotic, Fractal, and Nonlinear Signal Processing

R. A. Katz, Editor

American Institute of Physics Press, Woodbury, NY, 1996. xxi+858 pp. Price $165.00.

An international cadre of experts in the science of nonlinear systems gathered amid the warm salty sea breeze of July 1995 at the Seaman's Inne and Conference Center in Mystic, CT, a beautiful historic seaport known for its nineteenth century whaling industry. The participants were attracted to this conference by the promise of stimulating conversations about their most recent work. They were not disappointed. The Office of Naval Research, the Naval Undersea Warfare Center, and the International Science Foundation, sponsors for this the third in a series of biannual meetings, should be praised for having tripled the size of the meeting despite the dwindling defense spending for science and technology. In addition, Dr. Katz and his committee ably organized and coordinated the conference. The conference proceedings were recorded and published by the American Institute of Physics (AIP) and are the subject of this review.

The conference proceedings contain a table of contents, a forward, a preface, ten sections with one subsection, a list of participants, and an author index. Dr. Katz states in the forward that the book "...is intended for use by linear and nonlinear researchers alike as a seed for bed expansion of one's own ideas and new discovery." Clearly, the book is intended for those workers in the field who want to remain current in nonlinear science. In the preface, Drs. Arbarbanel and Moon point out the need for further support in nonlinear science especially in the areas of "anti-control" for chaotic systems, spatio-temporal problems, and complex adaptive systems. Nonlinear system "anti-control" is defined by them as the maintenance of a system in one chaotic state or transitioning through various chaotic states, thereby providing the capability of rapid response to changing inputs. This is a type of system adaptation for maintaining diversity.

It is evident from a look at the table of contents that nonlinear methods are now being applied in a wide range of disciplines. The proceedings include contributions in mathematical frontiers, predictability and control of chaos, detection and classification, applications in acoustics, advanced applied signal processing methods, stochastic resonance, machinery diagnostics, turbulence, geophysics, medicine, and new methods for modeling nonlinear systems.

Two areas of research new to this meeting are stochastic resonance and medicine. Stochastic resonance occurs in a Langevin system consisting of a multistable differentiable potential driven by periodic input signals with additive noise. In one application, stochastic resonance is used to enhance signal detection. The work reported at this meeting includes stochastic resonance in active and passive Chua circuits, the application of residence time in the characterization of hidden periodicities, and signal to noise improvement methods by the introduction of chaotic resonance that replaces stochastic noise with chaotic noise. These resonance techniques are used to model biological sensors and neurons. In the area of nonlinear processes in medicine, workers are applying theory to cardiology, hormone secretion, lymphatic systems, and biphonation in speech. The majority of the work reported here is time series analysis with the intent to identify features in the signal that indicate pathological behavior.

Of particular interest to the acoustic community is the subsection on applications in acoustics. The topics include signal detection and classification of marine mammal transients, acoustic cavitation, long-range underwater acoustic propagation using ray theory, low-frequency active target characterization using hidden Markov models, active noise reduction using a cooperative source to estimate chaotic and stochastic components using radial basis functions, and multifractal analysis of speech signals. A definite highlight of the meeting was high-speed holographic cinematography of acoustic cavitation by Drs. Lauterborn, Parlitz, Holzfuss, Billo, and Akhatov. The images show clouds of bubbles resonating in a driving sound field while moving along arborescent filaments. The period doubling of the filamentary structure was shown by holograms taken at the astounding rate of 69,300 holograms per second!

Some of the more significant applications from the other sections include Bayesian detection and classification of chaotic systems, traffic flow, kicked physical pendulum, cylindrical journal bearings, high-speed compressor, Earth's ozone field, sea surface temperature's distribution, oceanic jets, and cryptographic systems. Turbulence applications included wall pressure and shear fluctuations, transient pipe flow, Rayleigh–Bénard convection, and spontaneous bursting.

This book fully documents the meeting transactions, and reveals the progress of nonlinear signal processing of time series data from feature extraction and characterization by the calculation of parameters such as Lyapunov exponents, embedding dimension, fractal dimension, Hurst exponents, and entropy to system identification, estimation, and control. Some of these methods are very successful. However, an interesting cautionary note to nonlinear processing is raised by Drs. Anosov, Butkovskii, Kravtsov, and Surovyatkina. The caution is based on the nonstationary property of nonlinear systems, a property they call finite predictability time horizons. They show the required methods to solve these inverse problems under the conditions of step-wise parameter changes for low-dimensional systems.

Hopefully, funding will permit further work. For even though progress has been impressive as exemplified by this conference's proceedings, there is still much to be done in both the theory and the application of nonlinear science.

BRUCE J. BATES
*Naval Undersea Warfare Center Division, Newport
Newport, Rhode Island 02841*

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*
HARVEY H. HUBBARD, *325 Charleston Way, Newport News, Virginia 23606*
SAMUEL F. LYBARGER, *101 Oakwood Road, McMurray, Pennsylvania 15317*
D. LLOYD RICE, *11222 Flatiron Drive, Lafayette, Colorado 80026*
CARL J. ROSENBERG, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
WILLIAM THOMPSON, JR., *The Pennsylvania State University, University Park, Pennsylvania 16802*
ERIC E. UNGAR, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
ROBERT C. WAAG, *University of Rochester Medical Center, 601 Elmwood Avenue, Rochester, New York 14642*

5,525,041

43.25.Qp MOMENTUM TRANSFER PUMP

David Deak, New York, NY
11 June 1996 (Class 417/63); filed 14 July 1994

This pump adds momentum to fluid drawn into port **1** and exiting at port **5** through action of transducer **10** operating at ultrasonic frequency (e.g., MHz) with a sufficiently high amplitude to exert radiation pressure on
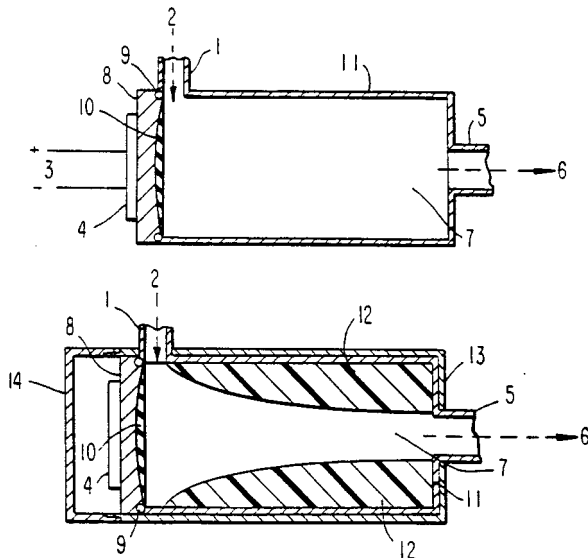
the medium passing through. Also shown is the addition of ''a well-defined tapered channel used to guide a focused ultrasound beam through the medium.''—DWM

5,469,407

43.30.Yj PARACHUTE ARRAY

R. I. Saunders, assignor to Lockheed Sanders, Incorporated
21 November 1995 (Class 367/173); filed 1 June 1981

A large sensor array is realized by attaching a plurality of hydrophones to the skin of a planar or volumetric parachute, or by affixing the hydrophones to strings positioned around such parachutes. As the parachute moves through the water, it expands and provides the necessary tension forces to position and maintain the hydrophone array in a predetermined geometric orientation.—WT

5,483,502

43.30.Yj METHOD AND APPARATUS FOR EMITTING HIGH POWER ACOUSTIC WAVES USING TRANSDUCERS

A. A. Scarpitta *et al.*, assignors to the French State
9 January 1996 (Class 367/158); filed in France 3 December 1993

The circular headmasses of a double-ended Tonpilz transducer (piezo-ceramic motor section) are fabricated from two materials. The central core of each headmass is made of a rigid material, such as steel, while the surrounding ring is of a lighter material, such as aluminum. Generally speaking, the coupling coefficient increases and the resonance frequency decreases as the percentage of steel increases. Hence the power handling capability, which varies as the product of frequency and the square of the coupling coefficient, is predicted to have a maximum for a specific steel percentage.—WT

5,600,133

43.35.Sx STRUCTURAL ANALYZER USING ACOUSTO-OPTIC SENSOR

William B. Spillman, Jr., assignor to Simmonds Precision Products, Incorporated
4 February 1997 (Class 250/227.14); filed 1 May 1995

For analyzing structures an elongated acoustic waveguide **12** is coupled to the structure **S** along its length to supply a pulse to embedded optical fiber **26**. Light transmitted through the fiber is modulated according
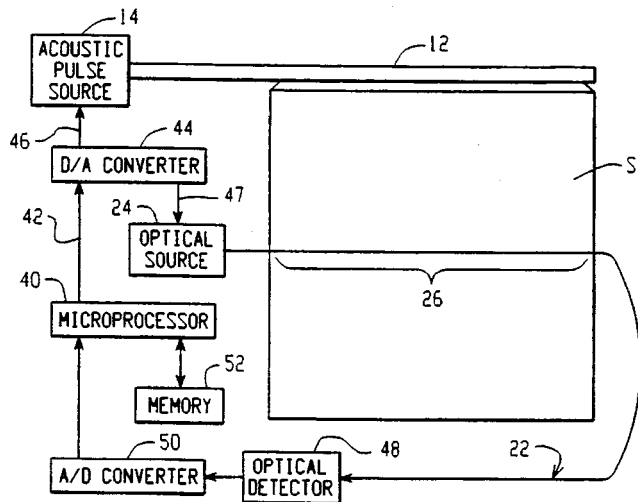
to the physical condition of that part of the structure through which the pulses propagate. Analysis of the transmitted acoustic signature can be used to locate and identify structural faults.—DWM

5,596,146

43.35.Ud PHOTOACOUSTIC MEASUREMENT OF UNBURNED CARBON IN FLY-ASH

David Waller and Robert C. Brown, assignors to Iowa State University Research Foundation, Incorporated
21 January 1997 (Class 73/590); filed 6 June 1994

A system for measuring the amount of unburned carbon in a sample of fly ash, using infrared photoacoustic absorption, directs radiation from source **10** modulated by modulator **12** toward a sample of fly ash in container **14** where ''the unburned carbon in the sample absorbs the radiation producing a thermal wave which propagates through the sample to generate
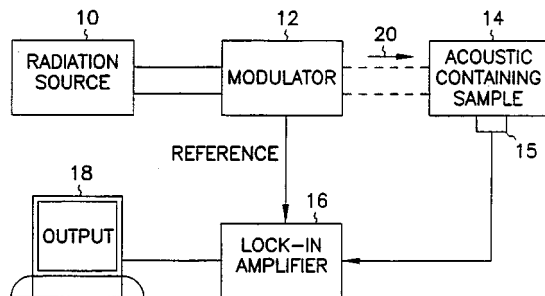
a minute acoustic wave at interfaces between the carbon particles and gas surrounding the particles.'' Microphone **15** detects the acoustic signals, and lock-in amplifier **16** improves the signal-to-noise ratio for computer analysis.—DWM

5,582,515

43.35.Ud ACOUSTICALLY PULSATING BURNER WITH INTEGRAL ADJUSTABLE SONDHAUSS THERMOACOUSTIC ELEMENTS

Frantisek L. Eisinger and Martin D. Bernstein, assignors to Foster Wheel Energy Corporation
10 December 1996 (Class 432/1); filed 4 May 1995

This patent describes an acoustically pulsating burner assembly ''which includes an elongated outer tube, a fuel supply means located within the tube, and utilizes one or more integral axially adjustable Sondhauss thermoacoustic tubular elements to generate pulsations within the burner. The acoustic pulsations are generated inside the axially movable closed-end tube element provided within the burner, and these vibrations promote more

efficient combustion of the fuel so as to reduce undesired emissions in the combustion products. If desired, multiple thermoacoustic tubular elements of different lengths can be advantageously used to generate pulsations of different frequencies within the burner assembly. The burner assembly is located in a furnace windbox, and can be adapted for combusting gas, oil or particulate fuels such as coal.''—DWM

5,594,293

43.35.Ud ELECTROACOUSTIC ENERGY CONVERTER FOR TRANSFORMATION BETWEEN THERMAL AND ELECTRICAL ENERGY

Victor Spivak, Haifa, Israel, assignor to Igor Gorlitsky
14 January 1997 (Class 310/339); filed in Israel 14 January 1992

The electroacoustic energy converter of this patent transforms thermal energy into electrical energy or vice versa. It consists of working chamber **1**, containing a working medium surrounding a piezoelectric transducer **300**, and connected hydraulically to a pair of oppositely disposed conical acous-

tical filters **4** with an inlet nozzle, and **5** with an outlet nozzle. Transducer **300** is wrapped in a glass fiber layer **300**′. The medium under pressure receives pulsations created during periodic ignition and combustion of the medium stream, generating an electrical output from the transducer. In the opposite function, electrical signals supplied to the transducer are converted to acoustical waves transferred through the working medium to the gaseous stream.—DWM
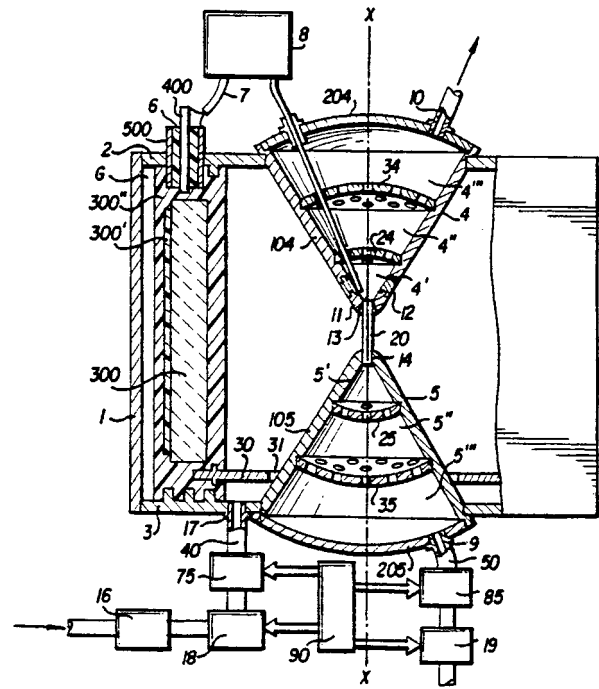
5,469,408

43.38.Dv HIGH RESOLUTION GEOPHONE

D. M. Woo, assignor to Shaw Industries Limited
21 November 1995 (Class 367/182); filed 20 July 1994

A cylindrically shaped moving-coil geophone, configured to an industry standard length of approximately 1.25 in., features a shorter, stronger magnet made from a rare earth material but longer than usual pole pieces. This results in a more uniform magnetic field in the annular space between the pole pieces and the surrounding case. As a consequence, the coils move through a more uniform field and the harmonic distortion in the output signal is said to be substantially reduced.—WT

5,523,524

43.38.Ja DOUBLE CHAMBER BASS REFLEX BOX

Jorg Prokisch and Markus Woldrich, assignors to Nokia Technology GmbH
4 June 1996 (Class 181/156); filed in Germany 24 December 1993

Chamber **14** and vent **24** appear to form some kind of auxiliary resonator to a conventional sealed/vented bandpass loudspeaker. In fact, they are
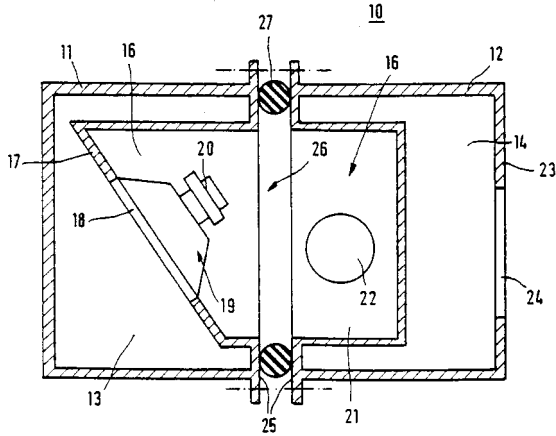
intended to function as an integrally molded port tube, allowing simple, inexpensive fabrication of a bandpass system.—GLA

5,590,214

43.38.Ja VERTICAL ARRAY TYPE SPEAKER SYSTEM

Hisatsugu Nakamura, Tokyo, Japan
31 December 1996 (Class 381/182); filed in Japan 12 November 1993

V-shaped baffles **1a**-**2a** and flared wings **7a**-**7b** form a simple horn driven by opposing arrays of small loudspeakers **3**. This is certainly a practical arrangement that does offer some advantages over more conventional
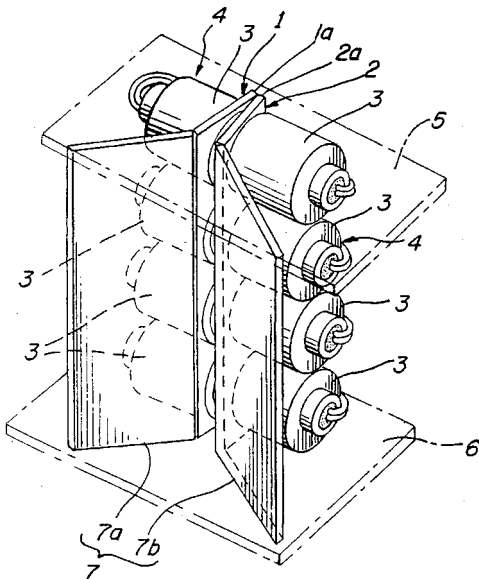
vertical arrays. Interestingly, the patent specifications insist that the baffles and wings be ''adjustable,'' but the intended meaning of the word is not entirely clear.—GLA

5,596,650

43.38.Ja EQUALIZING CIRCUIT FOR A LOUDSPEAKER SYSTEM

Stefan J. Hlibowicki, assignor to Audio Products International Corporation
21 January 1997 (Class 381/103); filed 29 April 1994

The patent describes a theoretician's loudspeaker equalizer. The goal is to extend the low-frequency response of a given loudspeaker system precisely and efficiently. A relatively simple circuit is described which cancels all the poles of the loudspeaker's known transfer function in the process of synthesizing a new transfer function.—GLA

5,548,650

43.38.Lc SPEAKER EXCURSION CONTROL SYSTEM

David L. Clark, assignor to Prince Corporation
20 August 1996 (Class 381/55); filed 18 October 1994

One or two state-variable filters are used in combination with a voltage-controlled amplifier to detect and limit signals that would otherwise cause excessive loudspeaker cone excursion.—GLA

5,574,791

43.38.Lc COMBINED DE-ESSER AND HIGH-FREQUENCY ENHANCER USING A SINGLE PAIR OF LEVEL DETECTORS

Robert A. Orban, assignor to AKG Acoustics, Incorporated
12 November 1996 (Class 381/98); filed 15 June 1994

As in an earlier DBX de-esser, two log-output level detectors receive an input signal, the first through a high-pass filter and the second through a bandpass filter. The difference between their outputs controls de-essing and, in this new device, program-adaptive equalization.—GLA

5,574,792

43.38.Lc VOLUME AND TONE CONTROL CIRCUIT FOR ACOUSTIC REPRODUCTION SETS

Fumiyasu Konno, assignor to Matsushita Electric Industrial Company
12 November 1996 (Class 381/103); filed in Japan 18 August 1993

To protect woofers from overload, commercial ''loudspeaker controllers'' may electronically boost low-frequency response at normal levels, then gradually reduce the boost as bass energy increases. In a simple home sound system loudness is controlled by a single knob. It follows that the operation of the bass control can be linked to that of the volume control ''... so that amplitude-versus-frequency characteristics of the output signal become flat when the movable contact of the volume control is set to maximum. The circuit will not distort a signal even at maximum volume.'' Well and good, but in this reviewer's experience consumers want to hear audible distortion when the volume control is halfway up; otherwise the system is judged not powerful enough.—GLA

5,589,830

43.38.Md STEREO AUDIO CODEC

Alfredo R. Linz *et al.*, assignors to Advanced Micro Devices, Incorporated
31 December 1996 (Class 341/110); filed 2 November 1994

This 61-page patent includes 58 illustrations, numerous tables, and highly technical text that will be easily understood only by digital circuit designers. However, it contains a wealth of information about up-to-date

CODEC circuitry, including D/A and A/D converters, digital noise reduction and filtering, control registers, data buses, etc.—GLA

5,555,310

43.38.Si STEREO VOICE TRANSMISSION APPARATUS, STEREO SIGNAL CODING/ DECODING APPARATUS, ECHO CANCELER, AND VOICE INPUT/OUTPUT APPARATUS TO WHICH THIS ECHO CANCELER IS APPLIED

Shigenobu Minami and Osamu Okada, assignors to Kabushiki Kaisha Toshiba
10 September 1996 (Class 381/17); filed in Japan 12 February 1993

The title of this Toshiba patent may be longer than that of Yamaha's best effort to date. The patent document itself is well written and includes more than a dozen informative diagrams. The application of the apparatus is two-channel audio (stereo) in combination with video teleconferencing via low-cost 64-kbps transmission links. Clever methods are described for maintaining stable stereo localization and including effective echo cancelation, all with a high degree of data compression.—GLA

5,600,718

43.38.Si APPARATUS AND METHOD FOR ADAPTIVELY PRECOMPENSATING FOR LOUDSPEAKER DISTORTIONS

Paul W. Dent et al., assignors to Ericsson, Incorporated
4 February 1997 (Class 379/406); filed 24 February 1995

A number of patents have been issued for digital circuitry that undistorts loudspeakers. This one is unusual in that the intended application is improved audio for cellular telephones. Rather than relying on fixed precompensating parameters, an adaptive filter is used. The handset microphone provides a feedback loop so that the filter can ''learn'' about such things as loudspeaker aging and environmental effects.—GLA

5,608,808

43.38.Si AUDIO-ADAPTED EYEGLASS RETAINER

Jean-Pierre M. da Silva, Salt Lake City, UT
4 March 1997 (Class 381/183); filed 24 October 1994

The patent shows an audio headset adapted to be attached to an eyeglass frame. Flexible, hollow, sleeve members carrying conductors for earphones are fitted over the ends of the eyeglass temples to supply the signals to earphones mounted on the temples. Conductors to a boom-type microphone may also be provided.—SFL

5,613,222

43.38.Si CELLULAR TELEPHONE HANDSET FOR HAND-FREE COMMUNICATION

Donald E. Guenther, assignor to The Creative Solutions Company
18 March 1997 (Class 455/89); filed 6 June 1994

The patent shows a hand-free device for receiving and transmitting speech signals from a cellular telephone. The unit is held in place by a hanger over the ear with a soft eartip projecting into the ear canal. A flexible hollow tube **15** carries sound into the ear canal from a cup strapped over the telephone receiver. Another hollow tube **116** conducts the user's voice from

a boom in front of the mouth to a cup strapped over the telephone microphone, or carries a cable that is connected to the microphone input terminals of the telephone. In the latter case, a small microphone is attached at the end of the boom.—SFL

5,579,396

43.38.Vk SURROUND SIGNAL PROCESSING APARATUS

Toshiyuki Iida et al., assignors to Victor Company of Japan
26 November 1996 (Class 381/18); filed in Japan 30 July 1993

Recent work in interaural cancellation techniques and pinnae function simulation has produced commercial stereo playback systems that create phantom surround sound sources from a single pair of loudspeakers. The apparatus at hand is intended to produce improved localization of virtual rear speakers. Apart from occasional computerese (''inputted'' and ''outputted''), the patent is clearly written and informative.—GLA

5,590,204

43.38.Vk DEVICE FOR REPRODUCING 2-CHANNEL SOUND FIELD AND METHOD THEREFOR

Hee-Soo Lee, assignor to Samsung Electronics Company
31 December 1996 (Class 381/24); filed in Republic of Korea 7 December 1991

The system makes use of two digital signal processors. One receives two-channel stereophonic signal, and transforms it into four channels. The second processor takes the resulting four channels and further modifies them for loudspeaker or headphone listening, ''... thereby reproducing 4-channel sound field feeling by 2 channels.''—GLA

5,596,645

43.38.Vk SOUND IMAGE LOCALIZATION CONTROL DEVICE FOR CONTROLLING SOUND IMAGE LOCALIZATION OF PLURAL SOUNDS INDEPENDENTLY OF EACH OTHER

Junichi Fujimori, assignor to Yamaha Corporation
21 January 1997 (Class 381/17); filed in Japan 30 March 1994

In two-channel stereophonic signal processing devices a monophonic source can be panned across the width of the sound stage by altering relative levels to left and right channels. More sophisticated processors introduce adjustable delay as well as gain. For a subjectively smooth pan, perhaps 25 delay stages must be provided for each source, gobbling up a lot of digital memory. The patent describes (at length) a clever method of turning the process inside out, requiring only one shift register or ring buffer per output channel.—GLA

5,586,182

43.40.Tm SHIP'S HULL VIBRATION DAMPER

Finn Orbeck, assignor to Orion Technology Limited
24 December 1996 (Class 114/65 R); filed in the United Kingdom 7 November 1991

In order to attenuate a ship's wave-induced vibrations at its fundamental mode, a tuned damper is mounted in the ship near an anti-node. A component of the ship, such as a cargo compartment or chain locker, serves as the damper's mass. This mass is supported from the hull via suitably selected springs and damping elements (shock absorbers), so that the natural frequency of this spring–mass system matches the ship's fundamental natural frequency reasonably well.—EEU

5,595,371

43.40.Tm VIBRATION ISOLATING SUPPORTER

Koumei Hukuda et al., assignors to Nippon Steel Corporation and Showa Electric Wire & Cable
21 January 1997 (Class 267/34); filed 8 April 1994

In this vibration isolation support several coil springs, arranged in mechanical parallel, support an equipment mounting plate or frame atop a base. Each coil spring sits on the base in a reservoir of a damping material (i.e., a liquid or gel). A cylindrical or other volume-displacing element, shorter than the springs, is located within each coil spring, parallel to its axis, and attached to the equipment mounting plate. As this mounting plate moves relative to the base, in any of the six rigid-body degrees of freedom, the springs provide restoring forces and the damping material is pumped through the spaces between the springs' coils, thereby providing damping.—EEU

5,596,140

43.40.Yq METHOD OF MONITORING VIBRATIONS IN VEHICLES

Christopher R. Boyce and Martin R. Haggett, assignors to Rover Group Limited
21 January 1997 (Class 73/105); filed in the United Kingdom 30 June 1994

This patent pertains to a method for monitoring an internal combustion engine in a vehicle, in order to detect misfiring. The vibrations of a liquid in a reservoir—e.g., of fuel in the fuel tank—are sensed and subjected to frequency analysis. Vibrations corresponding to sloshing of the liquid are discounted by the control unit, as are vibrations induced by the vehicle's passing over rough road surfaces. If it detects vibrations with frequency characteristics that correspond to engine misfires, the control unit sends an activation signal to an engine misfire indicator.—EEU

5,596,141

43.40.Yq TIRE RESONANCE FREQUENCY DETECTING SYSTEM HAVING INTER-WHEEL NOISE ELIMINATION AND METHOD FOR THE SAME

Yoshihiro Nishikawa et al., assignors to Nippondenso Company
21 January 1997 (Class 73/146.2); filed in Japan 4 August 1994

The system described in this patent is intended to provide a means for monitoring the air pressure in automobile tires via determination of their natural frequencies. Vibration signals from the tires on the two driving wheels are derived from wheel rotational velocity sensors. These signals, however, contain components that are unrelated to the tires' natural frequencies. As described in this patent, these contaminating signal components here are due to slight eccentricities of gears in the power transmission system or to gear vibrations, and thus the corresponding ''noise'' in the signals

from the two tires are correlated. This correlation is the basis for processing that is used to eliminate these noise components from the signals.—EEU

5,521,341

43.50.Gf SOUND ATTENUATOR

Reinhard Stief and Manfred Mattutat, assignors to Firma Carl Freudenberg
28 May 1996 (Class 181/295); filed in Germany 28 May 1993

This patent relates to the design of an element for absorbing sound in the atmosphere. The element consists of at least two chambers adjacent to each other in the direction that vibrations are introduced into the system. Chambers are enclosed by means of plastic films and metal foils and are hermetically sealed. The attenuator comprises a spring–mass system, its operations parameters being related to the configuration of the spring element, the chamber size, and its material thickness and stiffness.—HHH

5,587,563

43.50.Gf AIR HANDLING STRUCTURE FOR FAN INLET AND OUTLET

Muammer Yazici and Werner Richarz, assignors to Dipti Kr. Datta, Mississauga, Canada
24 December 1996 (Class 181/224); filed 16 June 1994

This patent relates to large air-handling units for buildings. Air duct silencers are described for both inlet and outlet configurations. Broadband noise reduction is provided by air flow splitters filled with absorbing mate-

rials, and by perforated sidewalls backed up with cloth-covered zero-erosion fiberglass materials. Hollow chamber resonators are provided for reducing the noise due to the peak blade passage frequency.—HHH
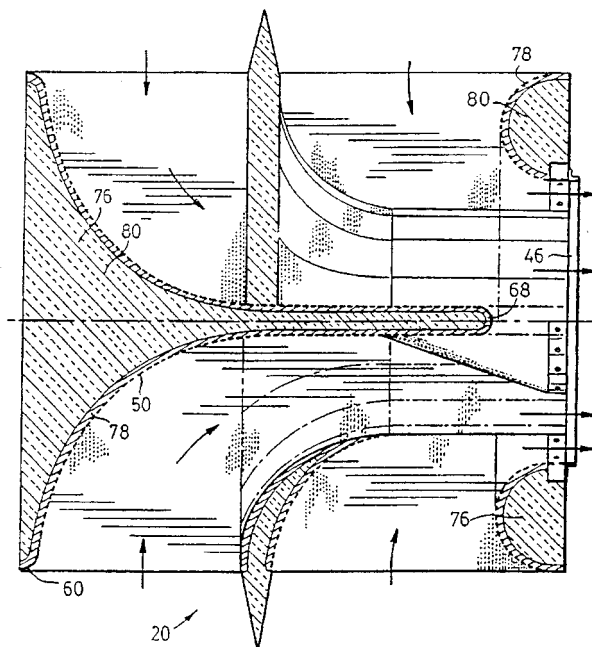
5,587,564

43.50.Gf NOISE DAMPER

Reinhard Stief and Gerhard Muller-Broll, assignors to Firma Carl Freudenberg
24 December 1996 (Class 181/295); filed in Germany 27 April 1994

This patent discloses noise dampers that provide broadband noise reduction over a frequency range 250–4000 Hz and are usable in damp locations and in clean rooms. They are comprised of two or more cup-shaped
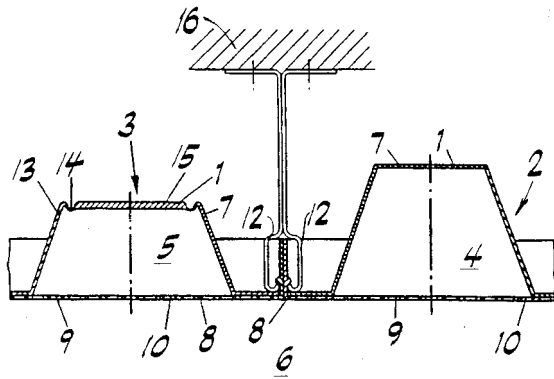
protrusions that open toward the noise source, each covered by an orifice plate having at least two openings and designed with different resonance frequencies.—HHH

5,588,810

43.50.Gf LOW NOISE REFRIGERANT COMPRESSOR

Michael A. Diflora *et al.*, assignors to Bristol Compressors, Incorporated
31 December 1996 (Class 417/312); filed 1 September 1995

This patent relates to the control of noise from refrigerant compressor units. An improved sound enclosure is described for enclosing the compressor unit with tongue and goove mating, with retention snaps, and with sound-absorbing material attached to inside surfaces. This shell-like enclosure is noted to be low weight, easy to assemble, and to provide easy access for servicing.—HHH

5,590,520

43.50.Gf METHOD OF ELIMINATING MACH WAVES FROM SUPERSONIC JETS

Dimitri Papamoschou, assignor to the University of California
7 January 1997 (Class 60/204); filed 5 May 1995

This patent relates to the elimination of Mach wave noise components from the exhausts of high-performance jet engines. This is accomplished by enclosing the primary jet stream with a slower moving jet stream such that
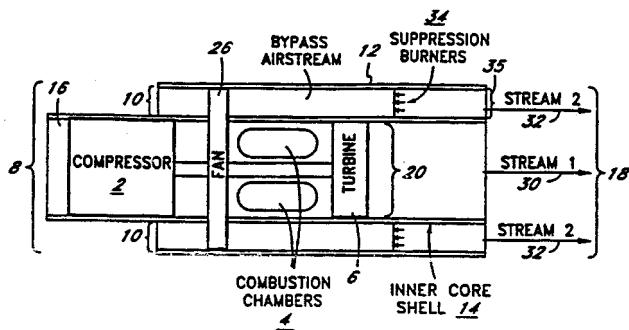
the convection speeds of the fast-moving primary jet eddies do not exceed the speed of sound in the ambient medium.—HHH

5,592,813

43.50.Gf HUSH KIT FOR JET ENGINE

Lee F. Webb, assignor to Avaero, Safety Harbor, FL
14 January 1997 (Class 60/226.2); filed 6 July 1995

This patent describes retrofit hardware for the reduction of noise from commercial aircraft turbo-fan engines. Included are provisions for an acous-

tically treated inlet and increased inlet guide vane spacing for fan noise reduction, and for improved mixing of the primary and secondary gas streams to reduce jet noise.—HHH

5,619,829

43.50.Gf SOUND INSULATING WALL AND METHOD OF INSTALLING THE SAME

Tokio Tan *et al.*, assignors to Nitto Boseki Company
15 April 1997 (Class 52/293.3); filed in Japan 26 January 1994

This is another kit of parts to make a highway noise barrier. The parts include a footing with special fittings, concrete panels, and steel threaded rods and nuts.—CJR

5,619,837

43.50.Gf CORRUGATED PANEL STRUCTURE

Fabricio N. DiSanto, Culver City, CA
15 April 1996 (Class 52/798.1); filed 26 July 1995

The patent focuses on a specially designed cap which fits atop a corrugated panel itself to provide stability. The structure could be used as a highway noise barrier.—CJR

5,594,803

43.50.Ki SILENCING APPARATUS OPERABLE TO REDUCE VEHICLE NOISE AT A TELEPHONE

Susumu Fujiwara *et al.*, assignors to Mitsubishi Denki Kabushiki Kaisha
14 January 1997 (Class 381/71); filed in Japan 11 March 1992

When a vehicle noise source (e.g., a train) approaches a telephone booth (e.g., on a station platform) the noise interferes with telephone communication. In this patent a series of distant microphones detects the approach, speed, and noise level of the approaching vehicle. When the level exceeds a predetermined threshold, the noise signal is amplified for inverse phase reproduction at the telephone booth for active noise cancellation to aid telephone communication. The noise signal may also be recorded for future use when the same type of vehicle approaches in the same way. Nearby microphones are also provided for vehicle noise pickup when the vehicle is adjacent to the telephone booth.—DWM

5,477,506

43.50.Yw IN-FLOW ACOUSTIC SENSOR

C. S. Allen, assignor to the United States of America
15 December 1995 (Class 367/140); filed 10 November 1993

A sensor for measuring the acoustic pressure in a fluid flowing over the sensor comprises a conventional transducer element (e.g., a capacitor or electret microphone) housed within a body of proper aerodynamic shape. The diaphragm of the sensor is located at an aft position relative to the nose of the body where unwanted noise is at a relatively low level.—WT

5,606,833

43.55.Ti WALL STRUCTURE

Leif A. Andersson, assignor to Isover Saint-Gobain, France
4 March 1997 (Class 52/506.01); filed in Sweden 27 August 1993

This patent describes a prefabricated wall panel comprising a rigid support sheet (like gypsum board), a soft insulating liner (like glass fiber insulation), and an inventive series of fastening brackets to attach the panels to floor, ceiling, and each other.—CJR

5,602,925

43.66.Ts HEARING AID WITH PROGRAMMABLE RESISTOR

Mead C. Killion, assignor to Etymotic Research, Incorporated
11 February 1997 (Class 381/68.4); filed 31 January 1995

The patent relates to a hearing aid with an integrated circuit having at least one programmable resistor for setting the audio response of the hearing aid. The programmable resistor has at least 14 discrete steps of programmed resistance. In one design the difference between successive resistance steps
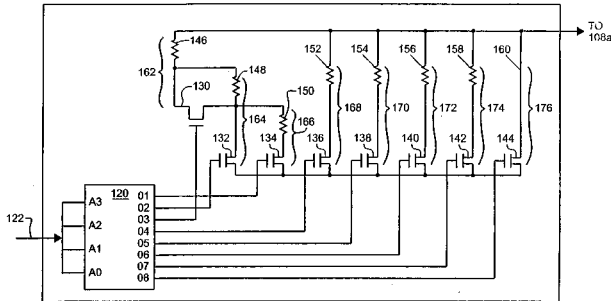
changes logarithmically. In a second version the difference in successive resistance steps changes logarithmically for the higher values, but linearly for the lower values of programmed resistance. Circuit details and tables of programmed resistor values are given.—SFL

5,604,786

43.66.Ts TELEPHONE WITH UNIFIED FEATURES FOR HEARING AND DEAF USERS

Robert M. Engelke et al., assignors to Ultratec, Incorporated
18 February 1997 (Class 379/52); filed 10 June 1994

The patent shows a combined TDD and telephone handset with ''in-

telligent mode'' switching to permit users to select either system or both, as desired.—SFL
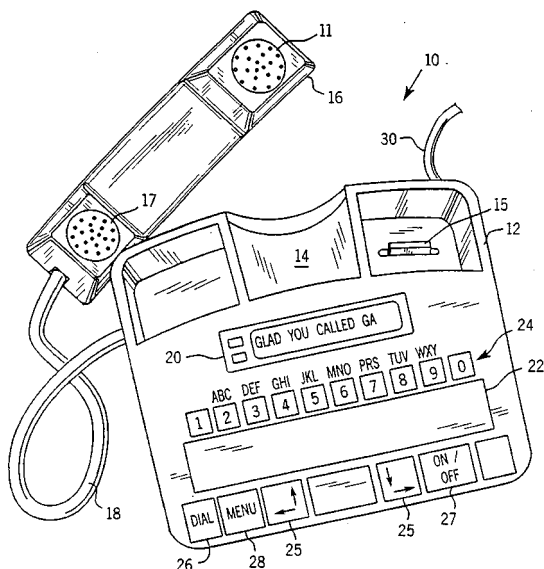
5,608,803

43.66.Ts PROGRAMMABLE DIGITAL HEARING AID

Neeraj Magotra et al., assignors to the University of New Mexico
4 March 1997 (Class 381/68.2); filed 17 May 1995

The hearing aid is provided with digital signal processor chips implemented as digital filters, having parameters that are established by erasable programmable read-only memories. Settings are based on a user's audiological characteristics. An additional digital filter changes frequency response, based on the characteristics of any background noise. Block circuitry is shown for a binaural system.—SFL

5,610,988

43.66.Ts HEARING AID SET

Katsunobu Miyahara, assignor to Sony Corporation
11 March 1997 (Class 381/68.4); filed in Japan 8 September 1993

The patent shows a behind-the-ear hearing aid which is totally controlled by a separate hand-held unit that sends encoded infrared signals to the aid. The controller sends signals to turn the aid on and off, to adjust
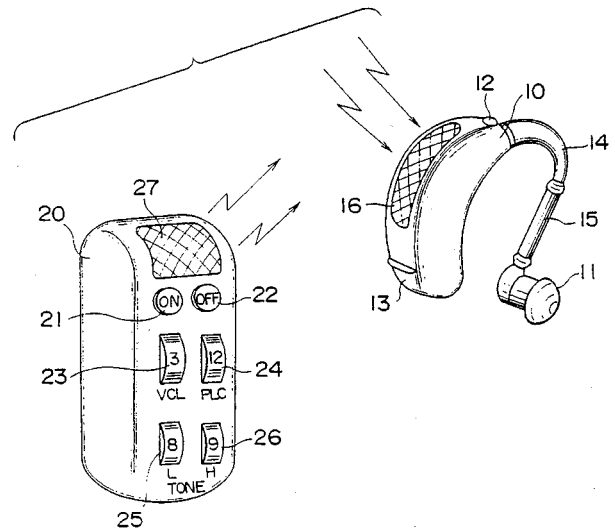
volume, to limit output, and to adjust low- and high-frequency response. Details of operation are given.—SFL

5,610,989

43.66.Ts COIL ASSEMBLIES

Richard J. Salvage et al., assignors to Knowles Electronics Company (UK)
11 March 1997 (Class 381/69); filed in the United Kingdom 21 December 1989

The patent relates to electrical coil assemblies, particularly for use in hearing aid transducers. The coils, usually of extremely fine enameled wire, are wound on a winding machine by fly winding. A flexible printed circuit carrier is cemented to the coil on the winder, and the end leads are welded to conductors on the carrier. The carrier and coil assembly can be accurately inserted directly into the transducer case without further need for attaching leads.—SFL

5,619,580

43.66.Ts HEARING AID COMPENSATING FOR ACOUSTIC FEEDBACK

Roy S. Hansen, assignor to GN Danavox A/S
8 April 1997 (Class 381/68.2); filed in Denmark 20 October 1992

The patent shows circuitry for a hearing aid having digital, electronic compensation for acoustic feedback. The digital compensation circuit includes a noise generator for the insertion of noise, and an adjustable digital filter for the adaptation of the feedback signal.—SFL

5,621,802

43.66.Ts APPARATUS FOR ELIMINATING OSCILLATION IN A HEARING AID BY USING PHASE EQUALIZATION

Ramesh Harjani and Rongtai Wang, assignors to the University of Minnesota
15 April 1997 (Class 381/68.2); filed 19 March 1995

To reduce acoustic feedback caused by leakage around the hearing aid shell, and through the vent provided to reduce the occlusion effect, phase equalization circuitry between the amplifier and the earphone is provided.
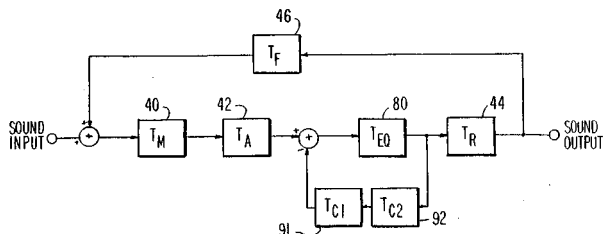
The phase equalization is accomplished by one or more first- or second-order filters that operate as an all-pass filter to provide a time delay without affecting the magnitude of the signal in the audio frequency range.—SFL

5,615,417

43.66.Vt EAR PROTECTOR

Claudia D. Jackson, New Orleans, LA
1 April 1997 (Class 2/209); filed 9 February 1996

The patent shows a protective pad for the outer ear that can be folded down when needed. The device is held in place by a forklike extension that fits over the pinna. Significant acoustical protection is not intended.—SFL

5,583,969

43.71.Ky SPEECH SIGNAL PROCESSING APPARATUS FOR AMPLIFYING AN INPUT SIGNAL BASED UPON CONSONANT FEATURES OF THE SIGNAL

Yoshiyuki Yoshizumi *et al.*, assignors to Technology Research Association of Medical and Welfare Apparatus
10 December 1996 (Class 395/2.63); filed in Japan 28 April 1992

This speech-enhancement system for hearing-impaired users operates on the premise that communication will be improved by increasing the signal gain during certain periods of certain consonant sounds. A plosive detector consists of two filter bands covering 100–900 Hz and 3700–5000
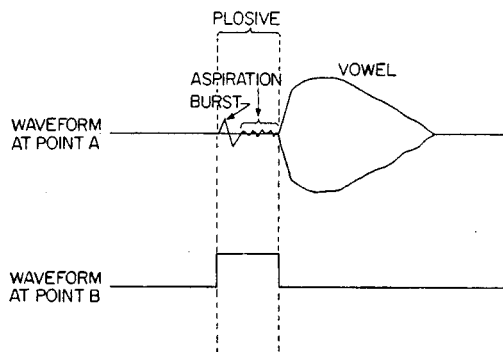
Hz and the computation of the ratio of high-band to low-band amplitude levels. The time duration for which the ratio exceeds a threshold indicates one of the voiceless plosives, /p/, /t/, or /k/. This indicator signal is used to control the gain of a delayed version of the input signal.—DLR

5,581,656

43.72.Ar METHODS FOR GENERATING THE VOICED PORTION OF SPEECH SIGNALS

John C. Hardwick and Jae S. Lim, assignors to Digital Voice Systems, Incorporated
3 December 1996 (Class 395/2.67); filed 20 September 1990

This improved speech analysis technique would be applicable to any of several types of speech coders in that it is primarily concerned with the detection of excitation events and the determination of the excitation period. Initial pitch estimates based on autocorrelation are refined by forward and backward smoothing and sample interval interpolation. A higher interpolation resolution for shorter periods maintains roughly constant percentage error. The voiced/unvoiced decision is improved by comparing the current period energy to that of one or more past periods. A combination of time and frequency domain methods is described for generating an output waveform.—DLR

5,583,968

43.72.Bs NOISE REDUCTION FOR SPEECH RECOGNITION

Michael Trompf, assignor to Alcatel N.V.
10 December 1996 (Class 395/2.41); filed in Germany 29 March 1993

A neural network is placed in the microphone signal pathway in the input to a speech recognizer. A mixed speech and noise signal is generated for training by mixing clean speech with an appropriate background noise signal. This allows the ''cascade'' network training algorithm to have access to both the noisy speech signal and the corresponding clean speech signal. A variation on the training procedure allows new nodes to be added to the network at any time without having to repeat the full training sequence.—DLR

5,586,215

43.72.Bs NEURAL NETWORK ACOUSTIC AND VISUAL SPEECH RECOGNITION SYSTEM

David G. Stork *et al.*, assignors to Ricoh Corporation
17 December 1996 (Class 395/2.41); filed 26 May 1992

This speech recognizer receives inputs from a traditional acoustic spectrum analysis as well as visual clues from a video camera. Acoustic and visual features drive a neural network which selects the most likely utterance in the manner of a template recognizer. The video processor locates key points on the facial display, providing lip movement measures as the

1282   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Reviews of Acoustical Patents   1282
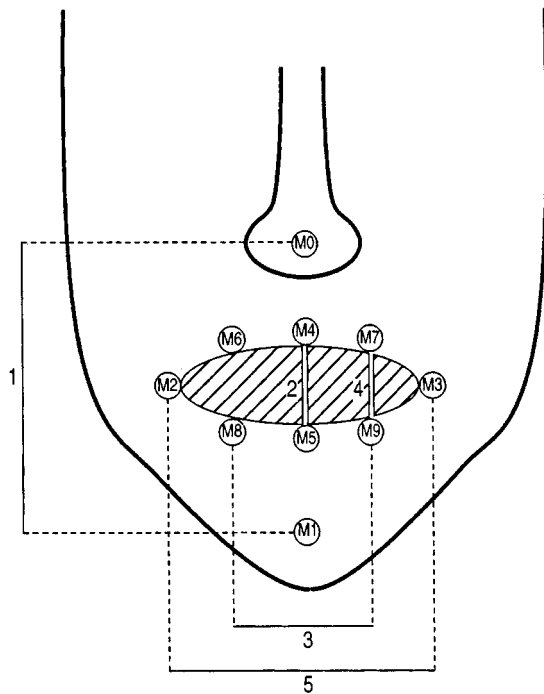
primary visual features. Error rate graphs suggest that training with both acoustic and visual features may not have a significantly lower error rate than acoustic features alone, but that the corresponding error rates are reached with roughly half the number of training passes.—DLR

5,581,652

43.72.Ew RECONSTRUCTION OF WIDEBAND SPEECH FROM NARROWBAND SPEECH USING CODEBOOKS

Masanobu Abe and Yuki Yoshida, assignors to Nippon Telegraph and Telephone Corporation
3 December 1996 (Class 395/2.31); filed in Japan 5 October 1992

This speech quality enhancer reconstructs a wideband signal from a telephone band signal covering typically 300–3400 Hz. Corresponding codebooks of spectral shapes are constructed from wideband and band-limited versions of a training data set. The narrow-band speech signal is then passed to a linear prediction analysis and a search of the narrow-band codebook. The corresponding wideband codebook entry is then used to synthesize a new speech signal at a higher sample rate.—DLR

5,579,431

43.72.Kb SPEECH DETECTION IN PRESENCE OF NOISE BY DETERMINING VARIANCE OVER TIME OF FREQUENCY BAND LIMITED ENERGY

Benjamin K. Reaves, assignor to Panasonic Technologies, Incorporated & Matsushita Electric Industrial Company
26 November 1996 (Class 395/2.23); filed 5 October 1992

This speech presence detector uses selected bands from a windowed Fourier transform of the speech signal to compute a frame-by-frame running variance. The energy level variance within a 250- to 3500-Hz band is computed for 12-ms frames with 50% overlap. Two thresholds mark beginning and confirming levels. If the variance is greater than the confirming level for at least 0.3 s, then the moment the beginning level was last exceeded marks
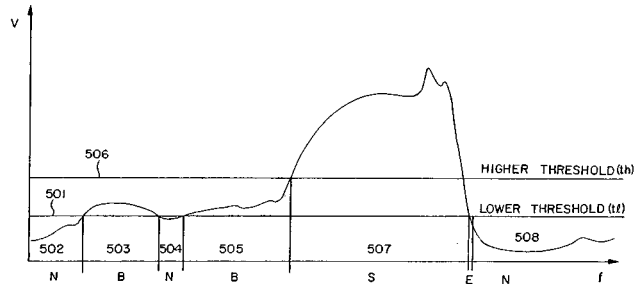
the beginning of speech. The end of speech is marked when the variance again falls below the beginning level. The use of variance is said to avoid problems related to absolute levels of gain and signal-to-noise ratio.—DLR

5,586,180

43.72.Kb METHOD OF AUTOMATIC SPEECH DIRECTION REVERSAL AND CIRCUIT CONFIGURATION FOR IMPLEMENTING THE METHOD

Achim Degenhardt and Thomas Henkel, assignors to Siemens Aktiengesellschaft
17 December 1996 (Class 379/389); filed in Germany 2 September 1993

This patent discloses a device for controlling the signal gains in both directions of a hands-free communication system, such as a speakerphone. A combination of speech versus noise detection, time-dependent attack and decay filtering and signal-to-noise ratio (SNR) tracking is employed to define three system states: neither party talking, one party talking, or both parties talking. The SNR values observed during these states control the gains to allow a more natural two-way interaction, such as when the listener attempts to interrupt the talking party.—DLR

5,579,433

43.72.Gy DIGITAL CODING OF SPEECH SIGNALS USING ANALYSIS FILTERING AND SYNTHESIS FILTERING

Kari J. Jarvinen, assignor to Nokia Mobile Phones, Limited
26 November 1996 (Class 395/228); filed in Finland 11 May 1992

This speech coder system uses a type of multipulse linear prediction coding. Following the short-term predictor computations, each stage of a series of pulse locators passes on the developing excitation pattern and the resulting residual to the next stage. After all stages of excitation pulse location, the final residual is vector quantized and the pulse positions and amplitudes are coded for transmission. This design would probably be rated as a potentially high-quality coder, but one having a fairly heavy computational load.—DLR

5,579,434

43.72.Gy SPEECH SIGNAL BANDWIDTH COMPRESSION AND EXPANSION APPARATUS, AND BANDWIDTH COMPRESSING SPEECH SIGNAL TRANSMISSION METHOD, AND REPRODUCING METHOD

Yasushi Kudo and Yoshiro Kokuryo, assignors to Hitachi Denshi Kabushiki Kaisha
26 November 1996 (Class 395/2.28); filed in Japan 6 December 1993

The patent describes a speech bandwidth compression system designed to transmit a high-quality speech signal over a narrow-band (analog-

modulated) radio signal. A linear prediction residual is computed from the speech signal and low-pass filtered to a cutoff at some fraction of the original sample rate. A low-frequency noise signal is added and the result used to synthesize a band-limited speech signal, which is downsampled, converted to analog, and transmitted over the channel. At the receiver, the full (wideband) LP coefficients are recovered and used to resynthesize the high-quality signal.—DLR

5,579,437

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

Bruce A. Fette *et al.*, assignors to Motorola, Incorporated
26 November 1996 (Class 395/2.71); filed 28 May 1993

This speech coding system consists of two separate branches; a combination of frequency domain techniques to process voiced portions of the input and time-domain techniques for the unvoiced portions. Frames of the linear prediction residual signal are autocorrelated to detect periodic excitation and to estimate its period. For unvoiced frames, rms levels in contiguous 1.5- ms time slots are vector quantized. For voiced frames, the moments of excitation are estimated, yielding fractional-sample excitation periods. The input signal is then rewindowed and a pitch-synchronous LP analysis is done. An additional ''phase correction'' is applied as the pitch-synchronous spectral vectors are selected for frame transmission.—DLR

5,581,651

43.72.Gy SPEECH SIGNAL DECODING APPARATUS AND METHOD THEREFOR

Toshiyuki Ishino and Akihiko Sugiyama, assignors to NEC Corporation
3 December 1996 (Class 395/2.14); filed in Japan 6 July 1993

This patent describes an improved method for handling transmission errors in a sub-band speech coder. At the receiving unit, the synthesizer is supplied with fill-in data to bridge the gap when a frame from the transmitter has been lost. The low-frequency bands are filled in by simply extending the band amplitude values from the last correctly received frame. For the high-frequency bands, average band amplitude values are maintained. On loss of a frame, white noise is adjusted to the average band amplitude to fill in the high bands.—DLR

5,583,963

43.72.Gy SYSTEM FOR PREDICTIVE CODING/ DECODING OF A DIGITAL SPEECH SIGNAL BY EMBEDDED-CODE ADAPTIVE TRANSFORM

Bruno Lozach, assignor to France Telecom
10 December 1996 (Class 395/2.28); filed in France 21 January 1993

This linear predictive speech coder uses an orthogonal transform to reduce the bitrate for the spectral envelope, freeing some of the bandwidth for an independent digital data stream. The front-end analysis, besides the usual long- and short-term prediction and perceptual weighting, includes

subtracting any previous frame energy remaining in the synthesis filter from the perceptually weighted excitation signal. The resulting short-term spectral vectors are transformed by an adaptive orthonormal transform similar to singular value decomposition.—DLR

5,579,436

43.72.Ne RECOGNITION UNIT MODEL TRAINING BASED ON COMPETING WORD AND WORD STRING MODELS

Wu Chou and Biing-Hwang Juang, assignors to Lucent Technologies, Incorporated
26 November 1996 (Class 395/2.53); filed 2 March 1992

Most speech recognizers, whether template, HMM, or neural network based, are trained by the accumulation of statistical characteristics from a limited amount of speech data. The most common training criterion is the error rate, that is, the percentage of the utterances in a testing speech data set which are not correctly recognized. According to this patent, a problem arises in that the error rate does not include information about the performance of near misses. A recognition loss function is described which takes into account the scoring of competing utterances similar to the correct one. The loss function is said to closely approximate the error rate, but provides a smooth error measure to replace the discrete error rate.—DLR

5,581,655

43.72.Ne METHOD FOR RECOGNIZING SPEECH USING LINGUISTICALLY-MOTIVATED HIDDEN MARKOV MODELS

Michael H. Cohen *et al.*, assignors to SRI International
3 December 1996 (Class 395/2.54); filed 21 January 1991

This hidden Markov model (HMM)–based recognizer employs groups of allophonically related models which cover dialectic variations of the sounds in an equivalence class. During training, a set of linguistic rules governs the activation of all of the HMMs in a class, allowing common feature characteristics to be trained for all allophones of the class. As allophone HMMs are accessed by specific pronunciations of a word, the rule set is automatically extended to include representations for the new class members.—DLR

5,583,961

43.72.Ne SPEAKER RECOGNITION USING SPECTRAL COEFFICIENTS NORMALIZED WITH RESPECT TO UNEQUAL FREQUENCY BANDS

Mark Pawlewski and Joseph G. Tang, assignors to British Telecommunications public limited company
10 December 1996 (Class 395/2.5); filed in the United Kingdom 25 March 1993

This patent relates to the use of cepstral coefficients for speaker recognition or verification. The coefficients would typically be nonlinearly distributed on the frequency axis, resulting in, for example, a mel-spaced frequency scale. Several transform methods are shown for computing the
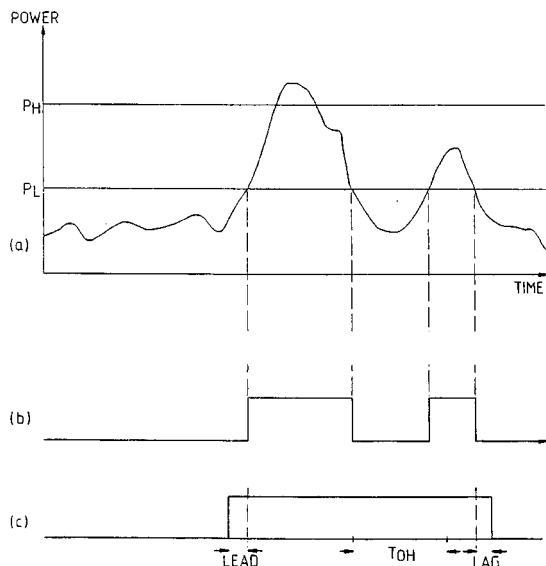
cepstral coefficients and mean removal from one or more of the cepstral values. It is not clear how this patent differs from other recent issues concerning the use of mel-scale cepstra for speech or speaker recognition.—DLR

5,583,965

43.72.Ne METHODS AND APPARATUS FOR TRAINING AND OPERATING VOICE RECOGNITION SYSTEMS

Peter Douma et al., assignors to Sony Corporation
10 December 1996 (Class 395/2.84); filed 12 September 1994

This brief patent presents a distributed training procedure for a speaker-dependent speech recognizer to help alleviate the burden of an intensive initial training session. Ideally, the user may proceed to use the system after a single training utterance. From that time on, the user may choose to provide additional training utterances at any time before using the system. It is a bit difficult to see how this strategy qualifies as nonobvious.—DLR

5,586,171

43.72.Ne SELECTION OF A VOICE RECOGNITION DATA BASE RESPONSIVE TO VIDEO DATA

Alex McAllister et al., assignors to Bell Atlantic Network Services, Incorporated
17 December 1996 (Class 379/67); filed 7 July 1994

Prior patents have addressed the fact that speech recognition performance, particularly in a telephone environment, may be improved by the use of any and all additional information on the caller. For example, the number
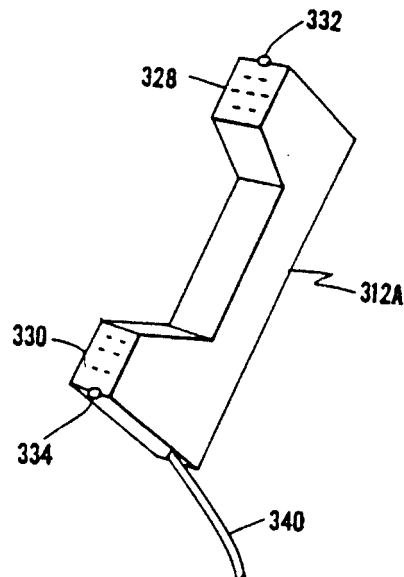
from a caller ID signal may yield demographic clues indicating a possible regional dialect. This patent considers additional information which may be gleaned from a video signal of the caller, possibly from a booth-mounted camera or headset-mounted LED pickup, where the caller need not be provided with a display for two-way video. A suggested application would be to improve the effectiveness of home incarceration.—DLR

5,585,579

43.75.De SOLID BODY CAPABLE OF VIBRATION AND/OR REFLECTION IN DEVICES AND INSTALLATIONS FOR GENERATING, RADIATING, DISTRIBUTING OR TRANSMITTING SOUND VIBRATIONS

George Ignatius, Hoefe, Germany
17 December 1996 (Class 84/192); filed in Germany 19 July 1983

The thickness of wood panels and braces in violins is typically rather uniform or tapered gradually. In this patent, thicknesses are purposely con-

toured or grooved in a complex manner to give the structure a more complex set of mechanical resonances.—DWM
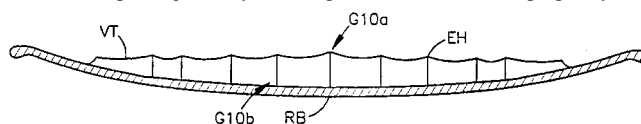
5,585,581

43.75.Hi GEL DRUMHEAD TRANSDUCING

Thomas P. Rogers, assignor to RTOM Corporation
17 December 1996 (Class 84/414); filed 16 January 1996

This amplified substitute for a conventional drum provides to the drummer a gel drumhead that is coupled to a mechanoacoustic transducer. ''The gel drumhead is constructed and arranged to present substantially the same resistive forces to drum sticks when struck presented by an acoustic drum correspondingly struck.''—DWM

5,610,350

43.75.Hi VARIABLE PITCH DRUM

Bradford W. H. Miller, Spencer, WV
11 March 1997 (Class 84/413); filed 5 May 1993

An African drum known as a ''Dondo'' has an hourglass-shaped body that can be squeezed to change drumhead tension that makes the instrument tunable during performance. This patent describes a foot-operated tuning
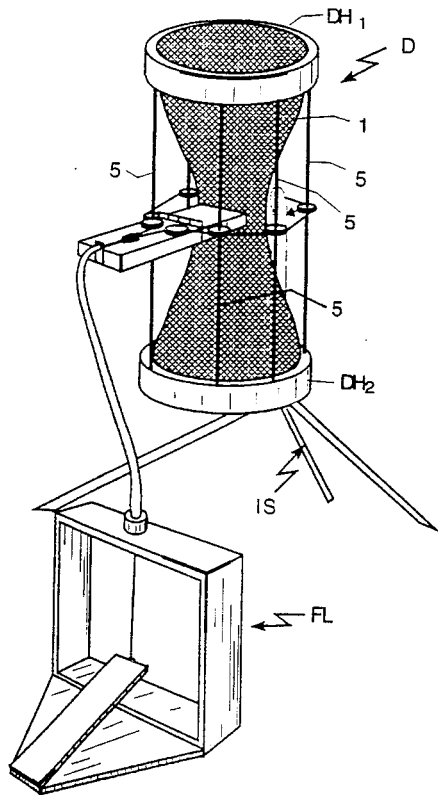
mechanism that controls a pulley system acting circumferentially on cables **5** connected to rings **DH₁** and **DH₂** to control the tension at both ends of the drum.—DWM

5,610,351

43.75.Hi MULTI-FACED BEATER FOR DRUMS

> **Mitsuo Yanagisawa, assignor to Pearl Musical Instrument Company**
> **11 March 1997 (Class 84/422.1); filed 29 February 1996**

The head of this drum beater has four convexly shaped faces, each differing from the others in either material hardness or shape or both. It can be rotated to provide a variety of drum sounds when struck by different faces.—DWM

5,500,486

43.75.Tv PHYSICAL MODEL MUSICAL TONE SYNTHESIS SYSTEM EMPLOYING FILTERED DELAY LOOP

> **Julius O. Smith III, assignor to Leland Stanford Junior University**
> **19 March 1996 (Class 84/622); filed 13 July 1993**

In a musical tone synthesizer of the ''physical modeling'' type, used for synthesizing a plucked or struck string tone, this patent describes a system in which the body filter can be eliminated by applying a suitable

---
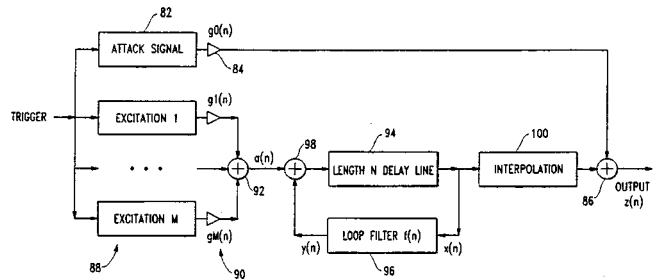
excitation signal **1** to a filtered delay loop having delay line **94** and loop filter **96**. This decreases the complexity of the system.—DWM

5,521,325

43.75.Tv DEVICE FOR SYNTHESIZING A MUSICAL TONE EMPLOYING RANDOM MODULATION OF A WAVE FORM SIGNAL

> **Chifumi Takeuchi et al., assignors to Yamaha Corporation**
> **28 May 1996 (Class 84/624); filed in Japan 22 March 1991**

For electronically or digitally produced musical tone signals to sound more like conventional musical instrument tones, a controlled random modulation is often required. This patent describes a digital system in which waveform generator **1** receiving information stored in memory **5** is used
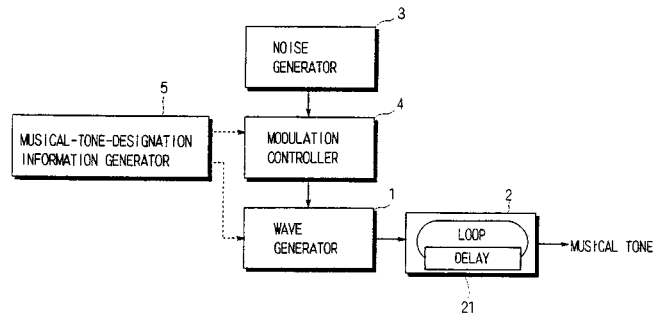
with loop section **2** containing delay circuits, filters, multipliers, and adders. A random noise generator **3** acts through modulation controller **4** to modulate the generated tone in a desirable manner with suitable filtering of the spectrum of the noise generator.—DWM

5,543,578

43.75.Tv RESIDUAL EXCITED WAVE GUIDE

> **Bryan J. Colvin, Sr. et al., assignors to Mediavision, Incorporated**
> **6 August 1996 (Class 84/622); filed 2 September 1993**

''Synthesizer models for emulating musical instruments can be improved using an analysis model that compares the output signal of the model to a recording of a desired sound and derives a residual signal that can be used to correct the model. When the original model is a good one, the residual signal is small and takes much less memory to store than is required for a sampled sound.''—DWM

5,610,354

43.75.Tv MUSICAL TONE SIGNAL SYNTHESIZER

> **Koyama Masahiro and Nishimoto Tetsuo, assignors to Yamaha Corporation**
> **11 March 1997 (Class 84/624); filed in Japan 12 January 1995**

In a digital electronic musical tone synthesizer this patent discloses a system for synthesizing the tone with a desired formant characteristic in a manner said to be simple. A first waveform signal is produced at a desired frequency. A second waveform signal is produced that is indicative of a

1286   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Reviews of Acoustical Patents   1286

windowing function, while a third waveform signal, synchronously controlled by the second signal, starts from a predetermined phase at each cycle of the second waveform signal. The first waveform is used to modulate the second waveform, and the second and third waveform signals are multiplied to produce the musical tone signal output.—DWM

5,504,270

43.75.Wx METHOD AND APPARATUS FOR DISSONANCE MODIFICATION OF AUDIO SIGNALS

William A. Sethares, Madison, WI
2 April 1996 (Class 84/645); filed 29 August 1994

The method and apparatus of this patent are intended to modify the dissonance of a musical sound signal by identifying the tonal partials by frequency and amplitude, then shifting the frequency of one or more of the partials and reproducing the more or less dissonant tone as desired. Referenced publications from our *Journal* include Plomp *et al.*, **38**, 548–560 (1965) and Kameoka *et al.*, **45**, 1451–1459, 1460–1469 (1969).—DWM

5,616,845

43.80.Pe ACOUSTIC SENSOR SYSTEM FOR INSECT DETECTION

Robert Hickling *et al.*, Oxford, MS
1 April 1997 (Class 73/584); filed 18 March 1994

The patent describes ''an apparatus for quickly, accurately and inexpensively inspecting agricultural commodities for insects. A basic concept of the present invention is to provide an insect detecting apparatus including isolation means for isolating the agricultural commodities from external noise and vibration; improved acoustic sensor means for detecting any sound from the agricultural commodities and for generating a signal in response to any sound so detected; and user recognizable output means for producing user recognizable output in response to signals generated by the acoustic sensor means ... (It) was developed initially to speed up inspection for pink bollworm in cotton bolls.''—DWM

5,602,924

43.80.Qf ELECTRONIC STETHOSCOPE

Jocelyn Durand *et al.*, assignors to Theratechnologies, Incorporated
11 February 1997 (Class 381/67); filed 9 December 1993

The patent shows an electronic stethoscope designed to minimize various types of noise. A first filter unit filters out the operator's tremor vibration and passes frequencies below 75 Hz; a second filter unit attenuates ambient noise and passes sound components of interest in the range 110–1300 Hz; and a third filter unit passes the sounds of mechanical heart valves. A level detector periodically disables the audio amplifier if the detected level exceeds a predetermined amplitude. A well-drawn block diagram of the circuitry is shown.—SFL

5,604,811

43.80.Qf STETHOSCOPE AND HEADSET SYSTEM

Kenneth M. McIntyre and Constance T. McIntyre, San Antonio, TX
18 February 1997 (Class 381/67); filed 29 September 1994

The patent describes a stethoscope system that is particularly adapted for use in monitoring a patient in an aircraft, and communicating the results to crew members through the aircraft's intercom system. The chest piece is cylindrical. It has a diaphragm at one end and also contains controls for adjusting the level of the received vibrations. The output of the chest piece goes to a switch that allows the signal to be routed to the aircraft's intercom

if desired. The headset is shown with a microphone on a boom from one earphone that allows communication through the intercom when desired.—SFL

5,610,987

43.80.Qf ACTIVE NOISE CONTROL STETHOSCOPE

Thomas R. Harley, assignor to University of Mississippi
11 March 1997 (Class 381/67); filed 12 March 1996

The stethoscope uses three sensors. One sensor is placed against the body area being auscultated. On the opposite side of this unit, a second sensor is used to pick up undesired ambient noise, which is cancelled out electronically. A third sensor is placed at a distance from the first to pick up
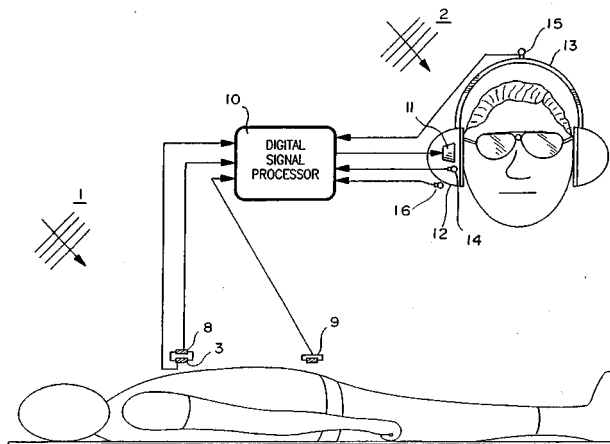
ambient noise for cancellation. The main sensor uses a piezoceramic pickup. Digital electronics are used to correlate the output of the sensors and deliver the processed signal to earphones.—SFL

5,616,890

43.80.Qf BINAURAL STETHOSCOPE ALLOWING SURROUNDING NOISES TO BE HEARD

Georges Boussignac, Antony, France
1 April 1997 (Class 181/131); filed in France 29 August 1995

The patent shows a chestpiece with two heads. The tubing from the usual binaurals is attached to a tube supporting the heads which continues beyond the heads a short distance, and that is open at the far end. The opening may be kept closed by a finger or may be left open to admit surrounding noise or speech. An alternate construction shows a small hole in the tube from the binaurals that can be closed by a finger or left open as desired.—SFL

5,595,182

43.80.Vj CARDIOVASCULAR MEASUREMENTS BY SOUND VELOCITY DILUTION

Nikolai M. Krivitski, assignor to Transonic Systems, Incorporated
21 January 1997 (Class 128/692); filed 7 June 1995

An indicator medium is injected into the bloodflow upstream of two ultrasonic transducers. The transducers detect the dilution through changes in sound speed measured using a transmission technique. The changes are used to determine hemodynamic parameters such as blood volume changes, protein concentration, or cardiac output.—RCW

1287   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Reviews of Acoustical Patents   1287

5,596,989

43.80.Vj ULTRASONIC PROBE

Katsuaki Morita, assignor to Olympus Optical Company
28 January 1997 (Class 128/660.1); filed in Japan 28 December
1993

In this probe, an ultrasonic transducer **1** is focused by a lens **2** onto a reflecting convex mirror **3(31)**. The result is a columnated beam perpendicular to the long axis of the probe. The beam is scanned circumferentially by
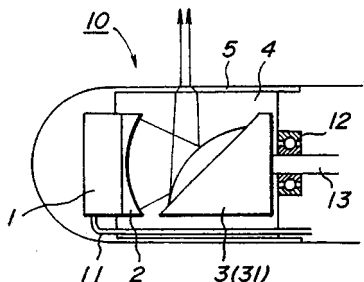
rotation of the convex mirror that is connected to a shaft **13** supported by a bearing **12**.—RCW

5,596,991

43.80.Vj CATHETER TYPE ULTRASOUND PROBE

Toshizumi Tanaka, assignor to Fuji Photo Optical Company
28 January 1997 (Class 128/662.06); filed in Japan 7 April 1994

The ultrasonic transducer in this probe is located at the end of an elongated flexible catheter and a rotatable support is located at the distal end. In the catheter is a flexible transmission shaft that is rotated by a controller mounted outside the catheter at its proximal end. Connections to the distal end are made through slip rings in the controller.—RCW

5,598,845

43.80.Vj ULTRASOUND TRANSDUCER DEVICE FOR CONTINUOUS IMAGING OF THE HEART AND OTHER BODY PARTS

P. Anthony Chandraratna *et al.*, assignors to Stellartech Research
Corporation
4 February 1997 (Class 128/662.03); filed 16 November 1995

This device contains a semispherical enclosure in which the transducer is mounted. The housing can be angled to scan the beam produced by the transducer. The assembly is fixed to the surface of the body by an adhesive patch.—RCW

Theory of the time-reversal process in solids[a),b)]

Carsten Draeger, Didier Cassereau, and Mathias Fink

Laboratoire Ondes et Acoustique, E.S.P.C.I., Université Paris VII, CNRS URA 1503, 10, rue Vauquelin, 75231 Paris Cedex 05, France

In this paper, a theoretical formulation is proposed to describe a time-reversal process in a solid medium with two propagation modes, longitudinal and transverse waves. A plane time-reversal mirror (TRM) is used, installed in a fluid which is in contact with the solid through a plane interface parallel to the TRM. The basic approach is similar to the case of a plane fluid–fluid interface [D. Cassereau and M. Fink, J. Acoust. Soc. Am. **96**, 3145–3154 (1994)]; it is extended to take into account the different wave types. It is shown that the TRM is able to recreate properly in time and space the reversed fields of the longitudinal wave and the vertical polarization SV of the transverse waves, but not the horizontal polarization SH. The focusing quality of the backpropagating waves is limited by their respective wavelengths, so the slower SV wave can be better focused. Additional, unwanted wavefronts are created in the solid, too, but they are of weak amplitude and they are not focused. Numerical simulations illustrate the dependence of the focal pattern on parameters like distance–source interface, size, and eccentricity of the mirror. © *1997 Acoustical Society of America.* [S0001-4966(97)04008-3]

PACS numbers: 43.10.Ln, 43.20.Gp, 43.20.Fn, 43.20.Bi [JEG]

LIST OF SYMBOLS

α, β, and c	velocities of longitudinal and transverse waves in the solid and of the pressure wave in the fluid
ρ_s, ρ	density of the solid and the fluid
$u(x,y,z,t)$	any function of time and space
$\tilde{u}(x,y,z,\omega)$	its Fourier transform over t
$U(k_x,k_y,z,t)$	its 2-D Fourier transform over x and y
$\tilde{U}(k_x,k_y,z,\omega)$	its 3-D Fourier transform over x, y, and t

INTRODUCTION

We have been working for several years on a technique to optimize acoustical focusing based on the principle of time-reversal invariance of the acoustic wave equation in lossless media.[1–4] In a few words, if $\varphi(\mathbf{r},t)$ is an acoustic field and thus is a solution of the wave equation, we know instantly that $\varphi(\mathbf{r},-t)$ is another solution and thus a possible acoustic field. In particular, if $\varphi(\mathbf{r},t)$ is a wave diverging from a point, $\varphi(\mathbf{r},-t)$ must be focused on this same location. This rises the idea of a focusing procedure in two steps: First, we install an acoustic source and measure its emitted field by means of a closed receiving surface around the medium. We suppose that each point of this surface is able to record the wave as a function of time, and then to reemit in a second step in order to generate the time-reversed solution. This time-reversed wave now backpropagates through the medium and focuses finally on the location of the initial source.

This concept of closed time-reversal cavity (CTRC) has been studied in a previous work.[3] However, a closed cavity is difficult to realize and often not useful for applications. If we replace the CTRC by a time-reversal mirror (TRM) of finite size, we can show that its performance can be comparable to the cavity, in spite of the loss of information.[4]

Our experimental time-reversal devices are mostly designed to work in a fluid environment. Nevertheless, a large field of applications prompts us to exploit the performances of the time reversal method in solid materials.[5] So we have to plunge the solid into water and consider the process via a solid–fluid interface. The case of two fluids separated by a plane interface has already been discussed in a previous work.[6] In this paper, we extend the theoretical approach to a homogeneous, isotropic solid bordered by a plane interface to a fluid and we discuss how the different modes of wave propagation are involved in the process.

We consider a pointlike source of elastic waves located inside a solid half space at the origin $x=y=z=0$ [Fig. 1(a)]. The plane solid–fluid interface is at $z=h>0$ and the TRM is located in the fluid at $z=Z>h$. The source emits a short pulse of longitudinal and transverse waves (or P and S waves) which are partially transmitted at the interface into the fluid and thus both converted into pressure waves. Unlike the previous paper[6] where we installed a second TRM to capture reflected waves, we now consider that all waves emitted or reflected into the negative z direction are lost. SH waves, in particular, are totally reflected and therefore it is impossible to apply the time-reversal process to this polarization of transverse waves. This results in a limitation of the TRM device. Only the P and SV wave components yield

---
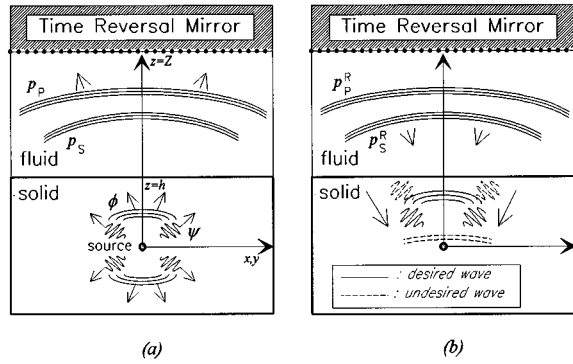
FIG. 1. Time-reversal behaves as a two-step process. (a): Emission of a short pulse of longitudinal and transverse waves by an active source in the solid, yielding two pressure-wavefronts in the fluid; Recording by the TRM (b): Re-emission of the time-reversed fields into the fluid by the TRM; Backpropagation of the two wavefronts in the fluid, yielding four wavefronts in the solid; The desired waves arrive simultaneously at the initial source location, the undesired ones arrive before and afterwards.

two wavefronts in the fluid, the first one created by the faster P, the second one by the slower SV waves.

The TRM records them both and is able to distinguish between the two wavefronts if their arrival time is sufficiently distinct. In this case, we can choose to throw back only the wavefront corresponding to the P waves, or the one of the SV waves, or both. When the backpropagating wavefronts arrive at the fluid–solid interface, each of them creates *two* wavefronts in the solid, one corresponding to the original type of wave, thus wanted, and one of the other type of wave, thus unwanted [Fig. 1(b)]. We will show that the unwanted SV wave (created by the returned P wavefront) as well as the unwanted P wave due to the returned SV wavefront only to yield a low-level noise, they are not focused and arrive at different times. On the contrary, the two wanted wavefronts focus at the same time and at the same place, i.e., the location of the initial source which now remains passive.

In Sec. I, we develop the theory to describe the time-reversal process for any type of acoustic source. The results obtained are illustrated in Sec. II using a special type of source. Numerical simulations are performed showing the focal pattern in several examples.

I. GENERAL THEORY

As SH waves, as well as the waves emitted by the source in the negative z direction, are of no interest in the time-reversal process, we describe the displacement vector $\mathbf{u}$ of the incident elastic field as a function of the potentials ϕ and ψ of the P and SV waves propagating in the positive z direction:[7]

$$\mathbf{u}(x,y,z,t) = \boldsymbol{\nabla}\phi(x,y,z,t) + \boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(0,0,\psi(x,y,z,t)). \tag{1}$$

In fact, there is no need that this initial field is due to a pointlike source. The conclusions drawn below will also be valid for extended sources. The only condition is the existence of the 2-D Fourier transform $\mathrm{FT}_{2\text{-D}}[..]$ over x and y in the frequency domain of the potentials, $\widetilde{\Phi}(k_x,k_y,z,\omega)$ and $\widetilde{\Psi}(k_x,k_y,z,\omega)$:

$$\phi(x,y,z,t) = \frac{1}{2\pi}\int d\omega\;\widetilde{\phi}(x,y,z,\omega)\exp(-j\omega t),$$

$$\widetilde{\phi}(x,y,z,\omega) = \left(\frac{1}{2\pi}\right)^2 \int\int dk_x\,dk_y\,\widetilde{\Phi}(k_x,k_y,z,\omega)$$

$$\times\exp(j(k_xx+k_yy)) = \mathrm{FT}_{2\text{-D}}^{-1}[\widetilde{\Phi}].$$

Due to the wave equation, the dependence on z can be written explicitly

$$\widetilde{\Phi}(k_x,k_y,z,\omega) = \widetilde{\Phi}(k_x,k_y,z=0,\omega)\exp(j\nu_a(k_r)z),$$

where ν_a is a function of $k_r = \sqrt{k_x^2+k_y^2}$ and the longitudinal wave speed α and is defined by

$$\nu_\alpha = \begin{cases} \sqrt{\omega^2/\alpha^2 - k_r^2}, & \text{if } k_r \leqslant \omega/\alpha \\ j\sqrt{k_r^2 - \omega^2/\alpha^2}, & \text{if } k_r > \omega/\alpha. \end{cases} \tag{2}$$

If ν_α is real, i.e., if $k_r \leqslant \omega/\alpha$, it can be considered as the z component k_z of the wave vector $\mathbf{k} = (k_x,k_y,k_z)$ of a propagative wave. If ν_α is imaginary, the wave is called inhomogeneous, its amplitude decreases exponentially with the depth in the z direction.

With similar definitions, we obtain for the shear-wave field:

$$\widetilde{\Psi}(k_x,k_y,z,\omega) = \widetilde{\Psi}(k_x,k_y,z=0,\omega)\exp(j\nu_\beta(k_r)z).$$

Each wave type in the solid generates at the interface a wavefront in the fluid. Describing the transmitted sound wave by its pressure field p, we can split it formally into the part created by the P wave p_P and the one produced by the SV wave p_s:

$$p(x,y,z,t) = p_P(x,y,z,t) + p_S(x,y,z,t), \quad \text{giving}$$

$$\widetilde{P}_P(k_x,k_y,z,\omega) = \widetilde{\Phi}(k_x,k_y,z=0,\omega)T_{\mathrm{Pf}}$$

$$\times\exp(j\nu_a h)\exp(j\nu_c(z-h)),$$

$$\widetilde{P}_S(k_x,k_y,z,\omega) = \widetilde{\Psi}(k_x,k_y,z=0,\omega)T_{\mathrm{Sf}}$$

$$\times\exp(j\nu_\beta h)\exp(j\nu_c(z-h)). \tag{3}$$

The transmission coefficients T_{Pf} and T_{Sf} are functions of k_r and ω, their exact expression is given in Appendix A; In (3), ν_c is defined similar to ν_α with sound-wave speed c.

At $z=Z$, the TRM records the arriving field. In previous papers,[4,5] we considered different cases depending on the type of baffle the TRM is mounted on. However, it turns out that the reemitted fields behave almost alike in all cases. So this time, we will leave out this distinction and consider only the easiest case, corresponding to a simple phase conjugation in the frequency domain.

We suppose that the mirror is infinitely large, i.e., it measures and emits at each point in the whole xy plane at $z=Z$, and we reintroduce its finite size by an obturation during the reemission step via an aperture function defined by

$$o(x,y) = \begin{cases} 1, & \text{at the active surface of the TRM} \\ 0, & \text{elsewhere.} \end{cases}$$

It has been shown[4] that the focal pattern converges to the one generated by an infinite TRM as its size increases.
We obtain for the reemitted fields at $z=Z$:

$$\tilde{p}_{P,S}^{R}(x,y,Z,\omega) = \left(\frac{1}{2\pi}\right)^2 o(x,y)$$

$$\times \int\int dk_x\, dk_y\; \tilde{P}_{P,S}^{*}(k_x,k_y,Z,\omega)$$

$$\times \exp(-j(k_x x + k_y y))$$

$$= \left(\frac{1}{2\pi}\right)^2 o(x,y)\,\mathrm{FT}_{2\text{-}D}[\tilde{P}_{P,S}^{*}].$$

Note that the former inverse Fourier transform is now a direct one.

To simplify forward mathematics, we omit the aperture function and hence calculate as if the TRM was of infinite size. However, in the simulations, we take into account its spatial limitations.

From (3) we obtain for the time-reversed wavefronts:

$$\tilde{P}_{P}^{R}(k_x,k_y,z,\omega) = \tilde{\Phi}^{*}(k_x,k_y,z=0,\omega)T_{\mathrm{Pf}}^{*}$$

$$\times \exp(-j\nu_\alpha^{*}h)\exp(-j\nu_c^{*}(Z-h))$$

$$\times \exp(-j\nu_c(z-Z)),$$

$$\tilde{P}_{S}^{R}(k_x,k_y,z,\omega) = \tilde{\Psi}^{*}(k_x,k_y,z=0,\omega)T_{\mathrm{Sf}}^{*}$$

$$\times \exp(-j\nu_\beta^{*}h)\exp(-j\nu_c^{*}(Z-h))$$

$$\times \exp(-j\nu_c(z-Z)).$$

Each of them generates two wavefronts in the solid. Hence we obtain two desired wave fields corresponding to the initial type of wave,

$$\tilde{\Phi}_{P}^{R}(k_x,k_y,z,\omega) = \tilde{\Phi}^{*}(k_x,k_y,z=0,\omega)\exp(-j\nu_\alpha z)$$

$$\times T_{\mathrm{Pf}}^{*}T_{\mathrm{fP}}\exp(-2\,\mathrm{Im}\,t\nu_\alpha h)$$

$$\times \exp(-2\,\mathrm{Im}\,\nu_c(Z-h)), \qquad (4)$$

$$\tilde{\Psi}_{S}^{R}(k_x,k_y,z,\omega) = \tilde{\Psi}^{*}(k_x,k_y,z=0,\omega)\exp(-j\nu_\beta z)$$

$$\times T_{\mathrm{Sf}}^{*}T_{\mathrm{fS}}\exp(-2\,\mathrm{Im}\,\nu_\beta h)$$

$$\times \exp(-2\,\mathrm{Im}\,\nu_c(Z-h)), \qquad (5)$$

and two undesired wavefronts, an SV wave created by a returned P wave and vice versa:

$$\tilde{\Phi}_{S}^{R}(k_x,k_y,z,\omega) = \tilde{\Psi}^{*}(k_x,k_y,z=0,\omega)\exp(-j\nu_\alpha z)$$

$$\times T_{\mathrm{Sf}}^{*}T_{\mathrm{fP}}\exp(j(\nu_\alpha-\nu_\beta^{*})h)$$

$$\times \exp(-2\,\mathrm{Im}\,\nu_c(Z-h)), \qquad (6)$$

$$\tilde{\Psi}_{P}^{R}(k_x,k_y,z,\omega) = \tilde{\Phi}^{*}(k_x,k_y,z=0,\omega)\exp(-j\nu_\beta z)$$

$$\times T_{\mathrm{Pf}}^{*}T_{\mathrm{fS}}\exp(j(\nu_\beta-\nu_\alpha^{*})h)$$

$$\times \exp(-2\,\mathrm{Im}\,\nu_c(Z-h)). \qquad (7)$$

With $\nu_\alpha=\nu_\alpha^{*}+2j\,\mathrm{Im}\,\nu_\alpha$ and $\mathrm{FT}_{2\text{-}D}[\tilde{\Phi}^{*}\exp(-j\nu_\alpha^{*})]$ $= \tilde{\phi}^{*}(x,y,z,\omega)$, the returned and desired P wave (4) may be written in the xy space as a convolution:

$$\tilde{\phi}_{P}^{R}(x,y,z,\omega) = \left(\frac{1}{2\pi}\right)^2 \tilde{\phi}^{*}(x,y,z,\omega)\;\underset{x,y}{*}\;\mathrm{FT}_{2\text{-}D}[T_{\mathrm{Pf}}^{*}T_{\mathrm{fP}}$$

$$\times \exp(-2\,\mathrm{Im}\,\nu_\alpha(h-z))$$

$$\times \exp(-2\,\mathrm{Im}\,\nu_c(Z-h))].$$

The first term, $\tilde{\phi}^{*}(x,y,z,\omega)$, corresponds exactly to the time-reversed P field $\phi(x,y,z,-t)$ we are interested in. But the quality of the reversed wave decreases by losses due to interface and propagation. These effects are carried by the second term, which convolves the first one. The spectrum of this term is mainly determined by its exponential functions. Their values are equal to 1 while ν_α and ν_c are real $(k_r\leq\omega/\alpha; k_r\leq\omega/c)$; for greater values of k_r, i.e., those representing the evanescent part of the spectrum, ν_α and ν_c are imaginary. The exponential functions give very small values, if the distances between interface h and observer z, or between interface and mirror Z-h, are greater than several wavelengths. The spectrum is cut off sharply at $k_r=\omega/\alpha$ or $k_r=\omega/c$. This truncation results in a focal spot width in the xy space which corresponds roughly to the greatest of the two wavelengths involved (normally, the one of the P wave). Numerical simulations presented in Sec. II show that the influence of the transmission–backtransmission coefficient $T_{\mathrm{Pf}}^{*}T_{\mathrm{fP}}$ on the spot size is less important. Of course, the amplitude of the focal spot is reduced but its shape is not altered very much; the information loss is almost exclusively due to propagation effects.

In the same way and with the same conclusions we write the returned, desired SV wave

$$\tilde{\psi}_{S}^{R}(x,y,z,\omega) = \left(\frac{1}{2\pi}\right)^2 \tilde{\psi}^{*}(x,y,z,\omega)\underset{x,y}{*}$$

$$\mathrm{FT}_{2\text{-}D}[T_{\mathrm{Sf}}^{*}T_{\mathrm{fS}}\exp(-2\,\mathrm{Im}\,\nu_\beta(h-z))$$

$$\times \exp(-2\,\mathrm{Im}\,\nu_c(Z-h))].$$

$\tilde{\psi}^{*}(x,y,z,\omega)$ corresponds to the time-reversed SV field $\psi(x,y,z,-t)$. This means in particular that, if the TRM reverses both fields together, they will focus at the same time at the same place, i.e., the initial source location. This is the proof of the time reversal mirrors capability of *spatial and temporal recompression*. Again, if the distances are great enough, the SV focal spot has approximately the size of the maximum of the two wavelengths and is thus smaller than the one of the P wave.

In the expression of both undesired wavefronts, (6) and (7), we find the term $\exp(j(\nu_\alpha-\nu_\beta^{*})h)$ or its complex conjugate. If h is not exceptionally small, this factor yields a phase that changes quickly with the angle of incidence, i.e., with k_x and k_y, and this inhibits focusing anywhere. (If h is very small, the undesired waves are actually focused, but this is due to the fact that the backpropagating wavefronts are already focused in the fluid.) Their amplitude is reduced in an important manner by the factors $T_{\mathrm{Sf}}^{*}T_{\mathrm{fP}}$ or $T_{\mathrm{Pf}}^{*}T_{\mathrm{fS}}$, where one term is small when the other is great and vice versa. Finally the undesired wavefronts do not arrive at the same time as the desired ones at the origin. The time-reversed wavefront corresponding to the P wave generates at the in-
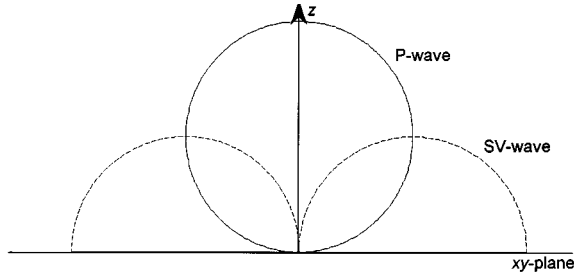
FIG. 2. Directivity pattern of P and SV waves created by a pointlike force parallel to the z axis. P waves are mainly emitted forward, SV waves mostly laterally.

terface two wavefronts in the solid: A desired one of P wave type and an undesired one of SV wave type which propagates more slowly than the first one and hence arrives later at the initial source position. In the same way, the time-reversed wavefront corresponding to the SV wave generates an undesired P wavefront in the solid which arrives sooner than the desired one.

So we find the main properties of the undesired waves: They arrive at the wrong time, they are not focused, and they are of low amplitude.

II. NUMERICAL RESULTS

In order to illustrate and quantify the previous results, we consider a particular ultrasonic source. Imagine a time-dependent body force inside the solid, only active at the origin and directed along the z axis:

$$\mathbf{f}(\mathbf{r},t)=(0,0,f(t)\delta(\mathbf{r})).$$

This type of force generates exclusively P and SV waves, but no SH waves. The resulting waves propagate in all directions, but we are only interested in the propagation in the positive z direction, i.e., towards the interface. As shown in Appendix B, we may describe the fields at $z>0$ by

$$\widetilde{\Phi}(k_x,k_y,z,\omega)=\frac{\widetilde{f}(\omega)}{2\rho_s\omega^2}\exp(j\nu_\alpha z)$$

and

$$\widetilde{\Psi}(k_x,k_y,z,\omega)=j\,\frac{\widetilde{f}(\omega)}{2\rho_s\omega^2}\,\frac{\exp(j\nu_\beta z)}{\nu_\beta}.$$

The corresponding directivity pattern is plotted in Fig. 2. P waves are emitted mostly in the forward direction, SV waves mainly in the xy plane.

To evaluate the focusing quality, we prefer to consider the displacement in the z direction u_z: It represents the direction of the initial excitation and it allows to compare directly P and S waves. It can be computed as

$$\widetilde{U}_{z\phi}^R(k_x,k_y,z,\omega)=\frac{\partial}{\partial z}\,\widetilde{\Phi}^R(k_x,k_y,z,\omega)$$

$$=-j\nu_\alpha\widetilde{\Phi}^R(k_x,k_y,z,\omega),$$

$$\widetilde{U}_{z\psi}^R(k_x,k_y,z,\omega)=(k_x^2+k_y^2)\widetilde{\Psi}^R(k_x,k_y,z,\omega)$$

$$=k_r^2\widetilde{\Psi}^R(k_x,k_y,z,\omega).$$
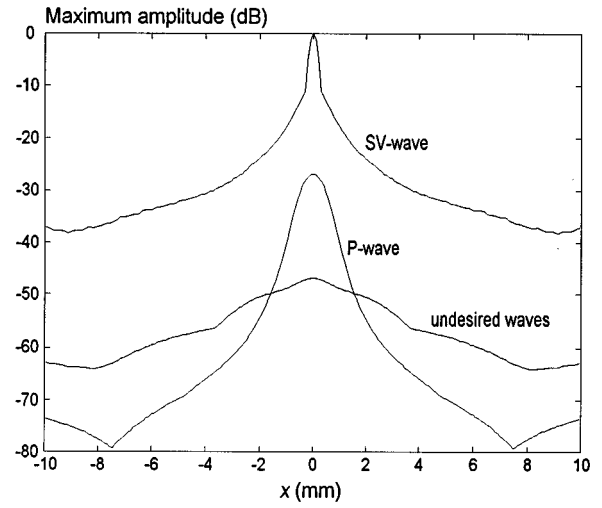
FIG. 3. Comparison of amplitude and focal spot size of P, SV, and undesired waves for a TRM of infinite size. SV waves have a higher amplitude and are better focused. The apparent weak focalization of the undesired waves vanishes completely for higher distances h to the interface.

With this formula, one may observe a curious property for the undesired waves: The focal patterns of their u_z component are identical, the spatial distributions in the xy plane are strictly the same. Of course, they do not arrive at the same time at $z=0$ and their u_x and u_y components are different.

We have developed a software to calculate the focal pattern of the backpropagating waves inside the solid. The waves and their propagation are first calculated in the k_xk_y space and in the frequency domain as indicated above, and the displacement u_z in real space and in the time domain is finally computed by fast Fourier transforms (FFTs). We obtain the focal pattern by displacing the observation point along the x axis in the plane $z=0$ and taking the maximum amplitude in time.

As an example corresponding to future experimental setups, we take the properties of duraluminum for the solid and water for the fluid:

$$\rho_s=2700\ \text{kg/m}^3,\quad \rho=1000\ \text{kg/m}^3;$$

$$\alpha=6300\ \text{m/s},\quad \beta=3150\ \text{m/s},\quad c=1500\ \text{m/s}.$$

The excitation function is a sinusoidal burst with a Gaussian envelope, a central frequency of 3.5 MHz and a relative -6-dB bandwidth of 100%. If not explicitly mentioned, the distance between source and interface is $h=5$ mm, and $Z=15$ mm. The spatial domain for x and y is in the range $[-25\ \text{mm},\ +25\ \text{mm}]$ with 512 steps of 0.098 mm. The wavelengths corresponding to the central frequency for all types of waves involved (P: 1.8 mm, S: 0.9 mm, water: 0.4 mm) are always large compared to this observation step. We use also 512 temporal points with a sampling frequency of 32 MHz.

First, we compare the z displacement of the returned P wave, S wave, and undesired waves in Fig. 3. The maximum amplitude is converted to decibel units into the range $[-80,\ 0\ \text{dB}]$ in order to have the greatest value of the three waves at 0 dB. The amplitude of the SV waves is shown to be higher than the P wave. This is due to several effects: The

1292   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Draeger *et al.*: Time-reversal in solids   1292
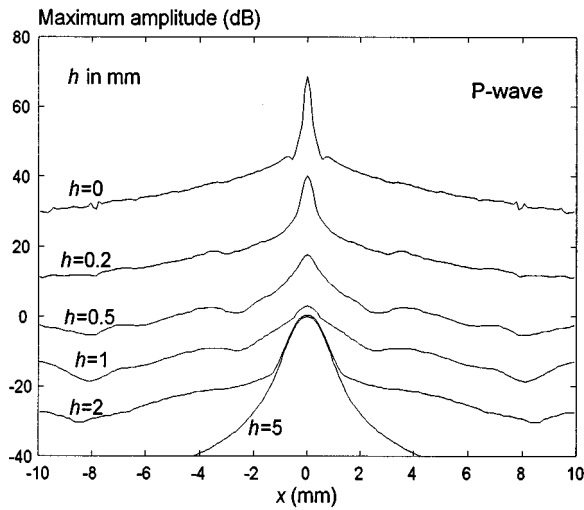
FIG. 4. Focal spot narrowing of the P wave by decreasing the distance h to the interface. For distances of $h = 5$ mm or higher, the evanescent part of the spectrum is insignificant, for distances of 0.5 mm or lower, the propagative part is negligible.

FIG. 5. P-wave focal spot dependence on mirror size: It widens slowly as the diameter D of the TRM decreases.

focal spot of the transverse waves is narrower and thus the maximum amplitude higher; Due to the higher impedance contrast, the transmission–backtransmission coefficient of the P waves is significantly smaller than the one of the SV waves; Finally, this particular type of source generates mainly transverse waves. The undesired waves are shown to be unfocused and of low amplitude.

The focal spot width does not result from the transmission coefficients, but from the decrease of the evanescent part of the spectrum during propagation. This fact is illustrated in Fig. 4, where we represent focal patterns of the P wave for different distances h of source and observer to the interface. Approaching the interface means reducing the propagation effects inside the solid, the truncation of the inhomogeneous waves as described in Sec. I is no more sharp. They even dominate the focal pattern for values of h of 0.5 mm or less (the propagating part does not depend on h). The evanescent waves permit a focal spot size smaller than the central wavelength of the P waves in the solid but, of course, not smaller than the pressure wavelength in the fluid because of the truncation of its spectrum during propagation in this medium.

However, sidelobe level rises for small values of h. The transmission–backtransmission coefficient of the propagating part of the spectrum can roughly be considered as constant and thus has little effect on the focal pattern. This is no longer true in the evanescent part. When this part dominates the spectrum, sidelobes are more important.

As in experimental setups we use mirrors of finite size, we now analyze the dependence of the focal pattern with respect to the size of the mirror. We introduce an aperture function $o(x,y)$ corresponding to a disk of diameter D centered on the z axis. In Figs. 5 and 6, we represent the influence of D on the P and SV waves, respectively. For values of $D > 30$ mm, we do not observe any significant difference to the case of an infinite mirror. The longitudinal focal spot enlarges continuously but slowly as D decreases. On the contrary, the transverse focalization holds its spot size, but
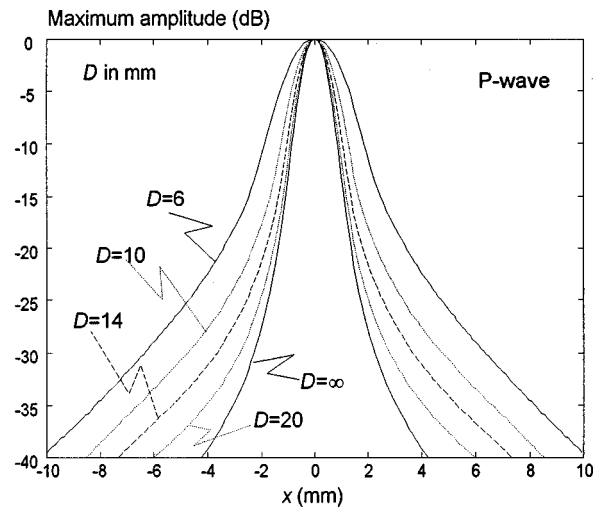
the sidelobe level rises quickly until complete breakdown of focusing.

SV waves are barely emitted in the forward direction and their transmission coefficient is low for small angles of incidence. So they yield only a weak signal in the region of the TRM near the z axis. One may be tempted to avoid this region by shifting the mirror laterally, and hence to time reverse only the signal arriving at one side of the center. In Figs. 7 and 8, we present focal patterns with a TRM of size $D = 14$ mm (a size giving an unsatisfactory result for SV waves), its center being located at various values X on the x axis. The focal pattern of the longitudinal waves suffers slightly from the shift, it remains centered at the initial source position for small values of X. On the contrary, for the transverse waves, we observe a breakdown of the focusing even for very small values of X. Thus a focusing enhancement by shifting the mirror from the center cannot be expected, we even have to pay attention to center TRM in
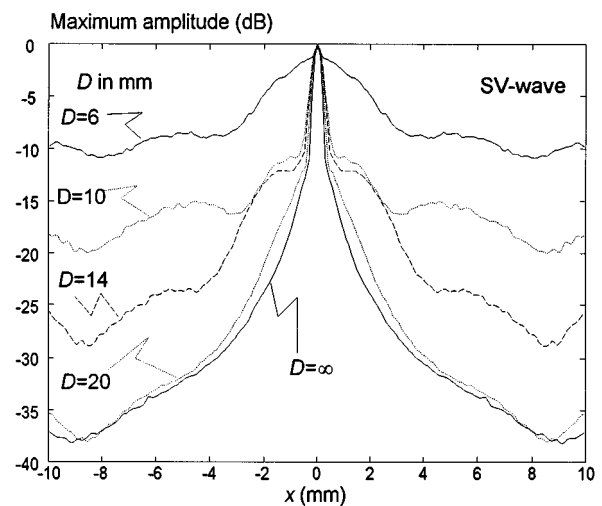
FIG. 6. SV-wave focal spot dependence on mirror size: Sidelobes rise quickly as the diameter of the mirror decreases.
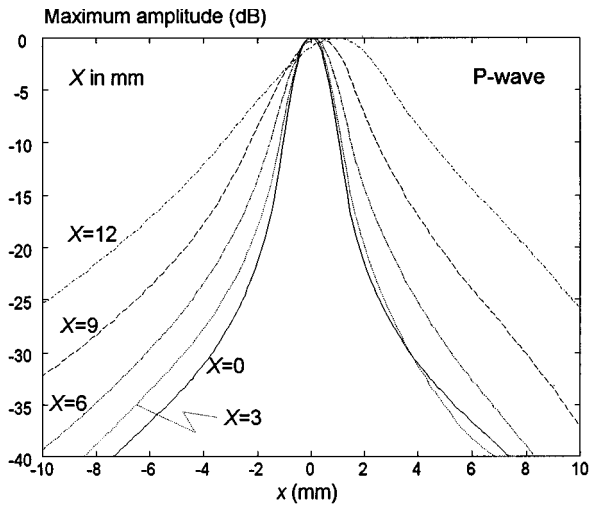
FIG. 7. Variation of the *P*-wave focal spot, if a mirror of diameter $D = 14$ mm is shifted by X from the center. The spot size increases slowly and loses its symmetry only for high values of X.

front of the source in order to minimize sidelobe levels.

III. CONCLUSIONS

In this paper, we give a theoretical approach to examine a time-reversal process in a solid medium with two propagation modes. The solid is bordered by a fluid where the time-reversal mirror works. To obtain refocusing quality, we take into account the different propagation modes in solid and fluid as well as the influence of their plane interface. It is shown that the TRM is able to perform a temporal and spatial recompression of the longitudinal (or *P*) waves and *SV* polarization of the transverse waves, but it cannot reverse the *SH* polarization. So the reversed *P* and *SV* waves arrive at the same time focused on the initial source location with a focal spot width corresponding approximately to their central wavelength. Thus the slower transverse waves are better focused. Furthermore, we predict the generation of two additional, unwanted, wavefronts. They do not harm the time-reversal process as they are of low amplitude, not focused, and of a different arrival time.
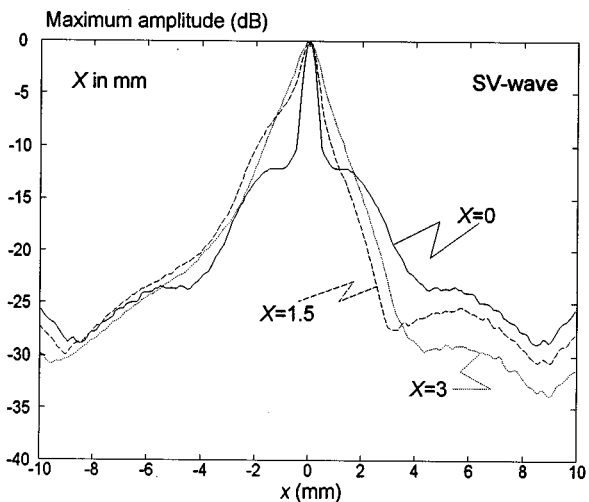
We have performed numerical simulations confirming these statements. They show furthermore that the main influence of the interface is limited to the loss of the *SH* waves, the focal patterns do not suffer in a significant way. *SV* waves are better focused, as predicted, but are shown to be also of higher amplitude than the *P* waves. However, they are more difficult to reverse: To obtain low sidelobe levels, they need a mirror of large aperture which has to be centered precisely in front of the source. Furthermore, we expect experimental difficulties for a pure *SV*-wave reversal: In solid samples of finite size, one may observe multiple reflections of longitudinal waves arriving at the mirror at the same time as the slower transverse wavefront. The TRM hence records a mixed signal which still can be reversed successfully, but the refocusing process does not yield the quality of a pure *SV* reversal.

First experiments using a laser source have been carried out. They confirm the main ideas developed in this paper and will be the subject of a following one.

APPENDIX A: TRANSMISSION COEFFICIENTS

In this Appendix, we are searching for the transmission coefficients for incident plane waves on a plane solid–fluid or fluid–solid interface. For simplicity, the interface is at $z = 0$ and $k_y = 0$, $k_x = k_r$. We consider three cases: (1) incident *P* wave, (2) incident *SV* wave, (3) incident sound wave. Giving to all incident waves the amplitude 1, the expressions of the total wave fields as defined in Sec. I [Eq. (1)] are

$$\text{1st case:} \begin{cases} \widetilde{\phi} = \exp(j(k_x x + \nu_\alpha z)) + R_{PP} \exp(j(k_x x - \nu_\alpha z)), \\ \widetilde{\psi} = R_{PS} \exp(j(k_x x - \nu_\beta z)), \\ \widetilde{p} = T_{Pf} \exp(j(k_x x + \nu_c z)). \end{cases}$$

$$\text{2nd case:} \begin{cases} \widetilde{\psi} = \exp(j(k_x x + \nu_\beta z)) + R_{SS} \exp(j(k_x x - \nu_\beta z)), \\ \widetilde{\phi} = R_{SP} \exp(j(k_x x - \nu_\alpha z)), \\ \widetilde{p} = T_{Sf} \exp(j(k_x x + \nu_c z)). \end{cases}$$

$$\text{3rd case:} \begin{cases} \widetilde{p} = \exp(j(k_x x - \nu_c z)) + R_{ff} \exp(j(k_x x + \nu_c z)), \\ \widetilde{\phi} = T_{fP} \exp(j(k_x x - \nu_\alpha z)), \\ \widetilde{\psi} = T_{fS} \exp(j(k_x x - \nu_\beta z)). \end{cases}$$

The fields are related to each other by boundary conditions at the interface $z = 0$ (see also Refs. 8 and 9):
(1) The *z* component of the displacement u_z is continuous:

$$\widetilde{u}_z = \frac{\partial}{\partial z} \widetilde{\phi} - \frac{\partial^2}{\partial x^2} \widetilde{\psi} = \frac{1}{\omega^2 \rho} \frac{\partial}{\partial z} \widetilde{p}.$$

(2) The vertical traction T_z is continuous:[7]

$$\widetilde{T}_z = \lambda \nabla^2 \widetilde{\phi} + 2\mu \frac{\partial^2}{\partial z^2} \widetilde{\phi} - 2\mu \frac{\partial^3}{\partial x^2 \, \partial z} \widetilde{\psi} = -\widetilde{p},$$

where λ and μ are the Lamé coefficients and related to the wave speeds by $\lambda + 2\mu = \rho_S \alpha^2$ and $\mu = \rho_S \beta^2$.
(3) The horizontal traction T_x vanishes:[7]

FIG. 8. Same as Fig. 7, for the *SV* wave. Focusing and symmetry break down even for small values of *X*.

1294   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997   Draeger *et al.*: Time-reversal in solids   1294

$$\widetilde{T}_x = 2\mu \frac{\partial^2}{\partial x\,\partial z}\,\widetilde{\phi} - \mu\frac{\partial}{\partial x}\left(\frac{\partial^2}{\partial x^2}\,\widetilde{\psi} - \frac{\partial^2}{\partial z^2}\,\widetilde{\psi}\right) = 0.$$

Writing these equations and eliminating the reflection coefficients for each case described above, we finally obtain the transmission coefficients:

$$T_{\text{Pf}} = \frac{2}{N}\,\rho_S\rho\omega^4(\omega^2 - 2\beta^2 k_r^2)\nu_\alpha,$$

$$T_{\text{Sf}} = \frac{4(-jk_r)}{N}\,\rho_S\rho\beta^2\omega^4 k_r\nu_\alpha\nu_\beta,$$

$$T_{\text{fP}} = \frac{2}{N}\,(\omega^2 - 2\beta^2 k_r^2)\nu_c,$$

$$T_{\text{fS}} = \frac{-4}{N(-jk_r)}\,\beta^2 k_r\nu_\alpha\nu_c,$$

with

$$N = 4\rho_S\beta^4 k_r^4\nu_c + 4\rho_S\beta^4 k_r^2\nu_\alpha\nu_\beta\nu_c - 4\rho_S\beta^2\omega^2 k_r^2\nu_c$$
$$+ \rho_S\omega^4\nu_c - \rho\omega^4\nu_\alpha.$$

APPENDIX B: INCIDENT FIELDS

In this Appendix, we are interested in the fields $\widetilde{\Phi}(k_x,k_y,z,\omega)$ and $\widetilde{\Psi}(k_x,k_y,z,\omega)$ resulting from a force located at the origin and directed along the z axis:

$$\mathbf{f}(\mathbf{r},t) = (0,0,f(t)\delta(\mathbf{r})).$$

An expression of the emitted fields $\phi_e(x,y,z,t)$ and $\psi_e(x,y,z,t)$ in the time domain and in the real space is given by Aki and Richards:[7]

$$\phi_e(x,y,z,t) = -\frac{1}{4\pi\rho_S}\left(\frac{\partial}{\partial z}\frac{1}{R}\right)\int_0^{R/\alpha}\tau f(t-\tau)d\tau,$$

$$\boldsymbol{\psi}_e(x,y,z,t) = \frac{1}{4\pi\rho_S}\left(\frac{\partial}{\partial y}\frac{1}{R}, -\frac{\partial}{\partial x}\frac{1}{R}, 0\right)$$
$$\times\int_0^{R/\beta}\tau f(t-\tau)d\tau,$$

where $R = \sqrt{x^2+y^2+z^2}$ represents the distance from the origin.

In the frequency domain we obtain

$$\widetilde{\phi}_e(x,y,z,\omega) = -\frac{\widetilde{f}(\omega)}{4\pi\rho_S}\left(\frac{\partial}{\partial z}\frac{1}{R}\right)\int_0^{R/\alpha}\tau\exp(j\omega\tau)d\tau$$

$$= -\frac{\widetilde{f}(\omega)}{4\pi\rho_S\omega^2}\frac{\partial}{\partial z}\left(\frac{\exp(j\omega/\alpha R)}{R} - \frac{1}{R}\right)$$

and

$$\widetilde{\boldsymbol{\psi}}_e(x,y,z,\omega) = \frac{\widetilde{f}(\omega)}{4\pi\rho_S\omega^2}\left(\frac{\partial}{\partial y}, -\frac{\partial}{\partial x}, 0\right)\left(\frac{\exp(j\omega/\beta R)}{R} - \frac{1}{R}\right)$$

$$= \nabla\times(0,0,\widetilde{\psi}_e).$$

The second term in each of the brackets, $1/R$, has no physical interest for the wave propagating in the positive z direc-

tion. In fact, the displacement $\mathbf{u}_{2\text{nd}}$ resulting from this term,

$$\mathbf{u}_{2\text{nd}} = \frac{\widetilde{f}(\omega)}{4\pi\rho_S\omega^2}\left[\nabla\left(\frac{\partial}{\partial z}\frac{1}{R}\right) + \nabla\times\nabla\times\left(0,0,-\frac{1}{R}\right)\right]$$

$$= \frac{\widetilde{f}(\omega)}{4\pi\rho_S\omega^2}\left(0,0,\nabla^2\frac{1}{R}\right),$$

vanishes everywhere but at the origin. Hence to describe the propagating fields, we set

$$\widetilde{\phi}(x,y,z,\omega) = -\frac{\widetilde{f}(\omega)}{4\pi\rho_S\omega^2}\frac{\partial}{\partial z}\left(\frac{\exp(j\omega/\alpha R)}{R}\right) \tag{B1}$$

and

$$\widetilde{\psi}(x,y,z,\omega) = \frac{\widetilde{f}(\omega)}{4\pi\rho_S\omega^2}\frac{\exp(j\omega/\beta R)}{R}. \tag{B2}$$

These two expressions show a term that corresponds to a spherical wave, whose 2-D Fourier transform can be written as follows (compare[7,10]):

$$\text{FT}_{2\text{-D}}\left[\frac{\exp(j\omega/\alpha R)}{R}\right] = 2\pi j\,\frac{\exp(j\nu_\alpha|z|)}{\nu_\alpha}. \tag{B3}$$

It finally results from (B1), (B2), and (B3) the following formula:

$$\widetilde{\Phi}(k_x,k_y,z,\omega) = -j\frac{\widetilde{f}(\omega)}{2\rho_S\omega^2}\frac{\partial}{\partial z}\left(\frac{\exp(j\nu_\alpha z)}{\nu_\alpha}\right)$$

$$= \frac{\widetilde{f}(\omega)}{2\rho_S\omega^2}\exp(j\nu_\alpha z),$$

$$\widetilde{\Psi}(k_x,k_y,z,\omega) = j\frac{\widetilde{f}(\omega)}{2\rho_S\omega^2}\frac{\exp(j\nu_\beta z)}{\nu_\beta},$$

valid in the range $z > 0$.

[1] M. Fink, ''Time-reversal of ultrasonic fields—Part I: Basic principles,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 555–566 (1992).
[2] F. Wu, J. L. Thomas, and M. Fink, ''Time-reversal of ultrasonic fields—Part II: Experimental results,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 567–578 (1992).
[3] D. Cassereau and M. Fink, ''Time-reversal of ultrasonic fields—Part III: Theory of the closed time-reversal cavity,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 579–592 (1992).
[4] D. Cassereau and M. Fink, ''Focusing with plane time-reversal mirrors: An efficient alternative to closed cavities,'' J. Acoust. Soc. Am. **94**, 2373–2386 (1993).
[5] N. Chakroun, F. Wu, and M. Fink, ''Ultrasonic nondestructive testing with time-reversal mirrors,'' Proc. IEEE Ultrason. Symp. Vol. 2, 809–814 (1992).
[6] D. Cassereau and M. Fink, ''Time-reversal focusing through a plane interface separating two fluids,'' J. Acoust. Soc. Am. **96**, 3145–3154 (1994).
[7] K. Aki and P. G. Richards, *Quantitative Seismology* (Freeman, New York, 1980), pp. 70–75, 135–138, 194–200, 215–219.
[8] E. Dieulesaint et D. Royer, *Ondes Élastiques dans les Solides* (Masson et Cie, Paris, 1974), pp. 166 ff.
[9] L. M. Brekhovskikh and O. A. Godin, *Acoustics of Layered Media I* (Springer-Verlag, Heidelberg, 1990), pp. 91–98.
[10] H. Bateman, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954).

1295   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Draeger *et al.*: Time-reversal in solids   1295

A study of materials at high temperature using miniaturized resonant tuning forks and noncontact capacitance transducers[a)]

David W. Schindel
Department of Physics, Queen's University, Kingston, Ontario K7L 3N6, Canada

David A. Hutchins
Department of Engineering, University of Warwick, Coventry CV4 7AL, England

Stuart T. Smith
Precision Engineering Laboratory, Department of Mechanical Engineering and Engineering Science, University of North Carolina at Charlotte, Charlotte, North Carolina 28223

A noncontact acoustical system has been developed which measures Young's modulus of solid materials at high temperature with $\leq 0.05\%$ accuracy. The system employs capacitive (or electrostatic) transducers to excite and detect vibrations of millimeter-sized resonant tuning forks, whose resonance frequencies alter with changing temperature and material properties. The use of tuning forks in their fundamental symmetric modes of vibration provide resonances of high Q's and eliminate irreversible frequency and drift effects that occur with other forms of resonator. The use of noncontact capacitive transducers reduces the damping and stresses that otherwise occur with contacting transducers, and allows the system to be simply and accurately modeled. Both single crystal silicon tuning forks, and those manufactured from hydrided and unhydrided Zr–2.5%Nb, were investigated at temperatures up to 700 °C. The measured responses of silicon forks confirmed the system accuracy, and suggested that single-crystal silicon be standardized as a calibration material for acoustical measurement systems. Results obtained with Zr–2.5%Nb tuning forks allowed an accurate quantitative analysis of the effects of hydride dissolution and precipitation on Young's modulus, and confirmed that Young's modulus of hydrided Zr–2.5%Nb decreases in proportion to free hydrogen concentration. This experimental system should prove valuable for accurate modeling of high-temperature material transformations in solids. © *1997 Acoustical Society of America.* [S0001-4966(97)05708-1]

PACS numbers: 43.10.Ln, 43.58.Hp, 43.35.Cg, 43.58.Vb, 43.38.Bs [SLE]

INTRODUCTION

Acoustic resonators are valuable tools for investigating material properties as a function of temperature. Their value stems from the general temperature dependence of a material's elastic constants, density, and internal friction, which effect changes in acoustic resonators as the temperature changes. In general, resonance frequencies are affected by changes in elastic constants and density, whereas strengths of resonances (or Q's) are affected by changes in internal friction. Suitable measurements of acoustic resonators as a function of temperature can therefore yield a wide range of information regarding the structure and strength of materials, structural transformations, thermodynamic and transport properties, etc.

There are many examples in the literature of the use of solid acoustic resonators for the study of temperature-dependent material properties. Many of the early techniques employed free torsional oscillations of thin wire specimens[1] or vibrating reeds[2–4] as a means to measure the temperature dependence of damping, elastic moduli, and dispersion. Results of such experiments have been instrumental in provid-

ing an understanding of internal friction processes and anelastic responses of solids.[5,6] One difficulty with these techniques is the appearance of frequency hysteresis and drift effects, which occur under conditions of cycling temperature due to unrepeatable strains that are induced in the resonators by their supporting clamps.

Early resonance techniques often monitored only one vibrational mode, but it has since been realized that increased accuracy results when numerous modes are observed. This multimode method, basically a form of acoustic resonance spectroscopy, gives improved accuracy since more than one modal frequency can be related to the same group of elastic moduli and sample dimensions. One example comes from standards research where all of the elastic constants of isotropic and anisotropic solids were obtained with accuracies of $\sim 0.1\%$ from measurements of the longitudinal, torsional, and flexural vibrations of solid bars.[7,8] The difficulties with this method arise mainly from the need for accurate placement of supports and the need to alter the support positions for each resonant mode of study.

In a recent variation of the multimode method, Migliori *et al.*[9,10] have studied small parallelepiped crystals of millimeter dimensions. Their method, termed rectangular parallelepiped resonance (or RPR), places two piezoelectric transducers on diametrically opposite corners of the sample, such

that one of the transducers drives the sample with a swept excitation frequency while the other measures the sample response. The resulting resonance spectrum is compared with theoretical predictions, and iterations are performed to give estimates for the elastic moduli. In the RPR technique, care must be taken to limit the piezotransducer contact forces to $<10^4$ dyn (equivalent to the load of a 10-g mass) so that unpredictable strains are not induced in the small samples.[10] Moreover, precise measurements of resonator dimensions and independent knowledge of thermal expansivity are required if accurate absolute values for elastic moduli are to be obtained.

At high temperatures the problems with the above methods are more severe, due to the added difficulties of oxidation, enhanced differential thermal expansion, creep of resonators, and the coupling of contacting transducers to heated samples. Furthermore, piezoelectric and magnetic-based transducers must remain below their Curie temperatures. For these reasons, the above-mentioned techniques are usually used at or below room temperature, although some high-temperature work has been reported.[11–17] The most successful high-temperature method has perhaps been the RPR technique, whereby elastic constants of single crystals have been determined at temperatures up to 1550 °C,[14–17] often through the use of piezotransducers mounted upon alumina temperature-isolation rods.[17] Unfortunately, these high-temperature RPR results have not yet been independently verified by a complementary technique of high accuracy, nor has the system employed a high-purity standardized sample for absolute calibration to date (as far as the authors are aware).

As a result of the many complications, it is necessary to further develop independent high-temperature resonance techniques and to establish standard samples that can be used to directly compare and calibrate the various methods. The need for calibration is supported by the fact that discrepancies are commonly observed in the literature in situations where the data of different methods are compared. Such developments are investigated in this paper, where noncontact transduction methods eliminate loading problems, while high-Q resonators with well-defined nodal points minimize the effects of hysteresis, creep, and damping at supports.

I. CHOICE OF RESONATOR

A wide range of resonators were initially investigated using capacitive and photoacoustic transduction schemes.[18–21] These included: vibrating reeds driven in flexural modes; thin circular disks clamped at their peripheries; axially supported solid spheres; and hemispherical shells (or bells). In short, all were found to exhibit either low Q's, or unrepeatable frequency responses with cyclic temperature variations. Furthermore, some of the resonators (such as the spheres) had the additional complication of being difficult to model theoretically. Based on these considerations, it was decided that the optimum resonator was a tuning fork, the two-dimensional equivalent of a bell, for it combined all of the advantages of the other resonators without the accompanying complications.[18]
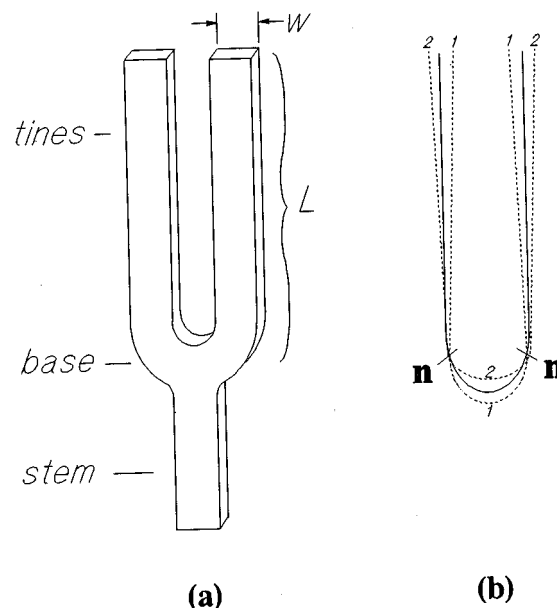
FIG. 1. Schematic diagrams describing (a) the components of a tuning fork and (b) its fundamental symmetric mode shape for three phases of the vibration.

The superior properties of a tuning fork have in fact been long appreciated. After all, the tuning fork played a significant role in the historical development of scientific measurements,[22] where it found extensive use as a standard pitch for experiments in acoustics,[23,24] allowed precision standards of frequency and time to be developed,[25] and provided accurate control over important scientific instruments through the development of the phonic motor.[25] The tuning fork has also proven useful in a wide variety of contemporary applications, such as watches,[26] paging systems,[27] thermometers for nuclear reactors,[28] angular rate sensors,[29] and the modulation of optical and molecular beams in ultrahigh vacuum systems.[30]

A tuning fork can be made in a variety of forms,[25,26,31,32] although the design of present interest is shown in Fig. 1(a). A U-shaped bar is manufactured from a continuous block of material, along with an integral support bar called the stem. Each of the two parallel tines of the U has a length L and a width w, and each is connected to the stem at what is here referred to as the base. During use, the bottom of the stem is rigidly fixed to a support by a suitable clamping mechanism.

The tuning fork is known to have a number of different types of resonance,[22,31,32] where each type generally consists of a series of resonant normal modes (i.e., a fundamental mode and its associated overtones). There are two series of cantilever-beamlike modes, for example, in which the whole structure bends back and forth, either in the plane of the page, or perpendicular to the plane of the page. Another type consists of a series of torsional-like modes in which one of the tines moves out of the page as the other moves into the page. Such cantilever and torsional modes, however, are not of interest for the present purpose, for they generate acoustic waves in the clamp at the bottom of the stem, and hence additional damping.

The type of vibration that is of interest is a series of

normal modes in which the tines vibrate in the plane of the page with mirror symmetry about the stem axis. In this case, the symmetrical vibrations of the tines generate equal and opposite strain fields in the base, which tend to cancel each other. Such cancellation results in a form of ''natural node'' at the base which isolates the vibrational energy of the tines from the stem, and provides a suitable clamping point with minimal energy loss.

The fundamental mode shape for the symmetric series of a tuning fork resonator (without a stem attached) is shown in Fig. 1(b). In this figure, the unstrained position of the tuning fork is indicated by the solid line and the extrema of the mode shape are given as the dotted lines 1 and 2;[25] note that the amplitude of the central antinode (i.e., between the nodes **n**) is greatly exaggerated in the diagram, as is the separation between the nodes. When the stem is added to the base, the nodes are brought even closer together, due to mass loading and stiffening of the base, and the amplitude of the central antinode further diminishes. This effectively results in a node at the base, although a small but finite vertical motion does remain. This finite motion leaks a small amount of energy into the stem, although certain designs can eliminate the loss entirely.[26]

There are two main advantages of the symmetric vibrations for the present purpose. First, since the tine vibrations are isolated from the stem through the natural node at the base, the resonances are of high Q and may be modeled by the same relationships that govern independently vibrating bars. Second, when the fork is operated in the fundamental symmetric mode, it acts as a quadrupole radiator of sound waves in the surrounding gas. As a quadrupole radiator is an inefficient source of sound, little power is radiated away from the fork through the surrounding fluid (only $\sim 3.5\%$ of the total[23]). Therefore apart from internal friction, all sources of energy loss for the symmetric modes are kept to a minimum.

II. THEORY OF TUNING FORKS AND THEIR TEMPERATURE DEPENDENCE

The symmetric modes of the tuning fork may be modeled by considering each tine as an independently vibrating cantilever beam which is fixed at the base and free at the tine tip.[25,32] For a thin tine of rectangular cross section, the resonance frequencies of flexural vibration f_n can then be written as[32]

$$f_n = \frac{\pi}{16\sqrt{3}} \frac{w}{L^2} \sqrt{\frac{E}{\rho}} \cdot k_n, \quad n=1,2,3,..., \tag{1}$$

where E is Young's modulus along the tine axis, ρ is the density, n is the mode number, and k_n is a mode-dependent multiplication factor given by the sequence

$$k_n = 1.194^2, 2.988^2, 5^2, 7^2, ..., (2n-1)^2. \tag{2}$$

Note that k_n generalizes to $(2n-1)^2$ for $n \geq 3$, whereas the fundamental ($n=1$) and first overtone ($n=2$) employ the respective values shown, namely 1.194^2 and 2.988^2. Equation (1) shows that the resonance frequencies of the tuning fork vary linearly with tine width and acoustic velocity

$(E/\rho)^{1/2}$ and as the inverse square of the tine length. Equation (2) shows that the overtones are not simple harmonics of the fundamental.

These equations apply specifically when w is small. Corrections for a tine of finite width (arising from effects of shear deformation and rotary inertia) can be estimated from a correction developed for vibrating reeds.[18] The approximate result shows that for the tuning fork dimensions and materials investigated below, and for the fundamental mode (with which we will be primarily concerned), this correction amounts to a $<2\%$ reduction in the absolute value of the fundamental frequency as predicted by Eq. (1). Other minor corrections would account for the curved portion of the tine near the base, and the effective nodal positions for improved estimates of tine length. Such corrections are usually evaluated through the use of finite-element analysis,[26] but are unnecessary in the present measurements as only the relative changes in resonance frequency with temperature are of interest.

The temperature dependence of the symmetric tuning fork modes can be obtained by allowing the resonator dimensions and material properties to be explicit functions of temperature. To do this, the tuning fork's dimensions and density can be written at any temperature T as

$$w(T) = w_0 \left[1 + \int_{T_o}^{T} \alpha_l(T) dT \right], \tag{3}$$

$$L(T) = L_0 \left[1 + \int_{T_o}^{T} \alpha_l(T) dT \right], \tag{4}$$

$$\rho(T) = \rho_0 \left[1 + \int_{T_o}^{T} \alpha_v(T) dT \right]^{-1}, \tag{5}$$

where w_0, L_o, and ρ_o are the respective values at room temperature T_o; $\alpha_l(T)$ is the linear (or one-dimensional) coefficient of thermal expansion; and $\alpha_v(T)$ is the volume coefficient of thermal expansion. If, in addition, Young's modulus $E(T)$ is represented by its room-temperature value E_o multiplied by a normalized temperature dependence $e(T)$, such that $E(T) = E_o \cdot e(T)$, then the frequencies of the tuning fork can be written with the aid of Eqs. (1) and (3)–(5) as

$$\frac{f_n(T)}{f_{n,o}} = \frac{[1 + \int_{T_o}^{T} \alpha_v(T) dT]^{1/2}}{[1 + \int_{T_o}^{T} \alpha_l(T) dT]} \cdot \sqrt{e(T)}. \tag{6}$$

Equation (6) shows that the temperature variation of the symmetric mode frequencies, when normalized to their room-temperature values $f_{n,o}$, do not depend upon the dimensions of the tuning fork or the mode number. Instead, all of the frequencies vary with temperature in the same way: according to the square root of the normalized Young's modulus multiplied by a function involving thermal expansion coefficients. The thermal expansion function, in fact, does not vary significantly from unity, such that the majority of the temperature variation of the tuning fork arises from the temperature dependence of Young's modulus alone. This can be seen for the case of aluminum, which has a large coefficient of thermal expansion compared with many solid

materials, $\alpha_l(T) = 23.6 \times 10^{-6} \ °C^{-1};^{33}$ yet even for a temperature change of 600 °C, the thermal expansion function of Eq. (6) deviates from one by less than 0.7%. Consequently, the normalized frequencies are relatively independent of the thermal expansion function. This does not mean that thermal expansion corrections can be ignored; only that the effects are small. Indeed, the experimental system to be described has accuracy and repeatability better than 0.7%, and so effects of thermal expansion must be carefully considered for each material of interest.

A simple rearrangement of Eq. (6) allows the normalized temperature dependence of Young's modulus $e(T)$ to be expressed in terms of tuning fork frequencies and thermal expansion coefficients as

$$e(T) = \frac{[1 + \int_{T_0}^{T} \alpha_l(T)dT]^2}{[1 + \int_{T_0}^{T} \alpha_\nu(T)dT]} \cdot \left[\frac{f_n(T)}{f_{n,0}}\right]^2. \tag{7}$$

This equation shows that, given the temperature dependence of thermal expansion coefficients, the temperature variation of Young's modulus can be obtained from measurements of tuning fork resonant frequencies. Moreover, since all of the symmetric mode frequencies vary with temperature in the same way, it is sufficient to monitor just one mode in order to determine $e(T)$. This being the case, the fundamental symmetric mode is concentrated upon below.

III. TUNING FORK FABRICATION AND CHOICE OF MATERIALS

Any further reduction of the above equations requires a specification of the dimensions and material properties of the resonator. In this paper two types of material are of interest: single crystal silicon, which has applications to calibration and thermometry; and a zirconium niobium alloy, which is of interest for studying material transformations.

A. Tuning forks of single-crystal silicon

Single-crystal silicon is one of the best possible choices for the resonator material, particularly when thermometry and system calibration are of interest. As an elemental single crystal, it eliminates unrepeatable frequency effects that can otherwise occur with cycling temperature due to grain boundary processes, species migration, phase transformations, etc. As such processes are also responsible for internal friction in solids, the use of a single crystal reduces damping and leads to a high Q resonator with a more repeatable frequency-temperature response curve. Furthermore, silicon is available with unparalleled purity and crystalline perfection, and its material properties are characterized to a high degree of precision.[34–38] This provides a readily available standard material which can be used for the comparison and calibration of measurement systems. Additional advantages of silicon include: high thermal diffusivity (which discourages temperature gradients in the resonator); and the lack of any structural transformations (providing a simple calibration for thermometry applications).
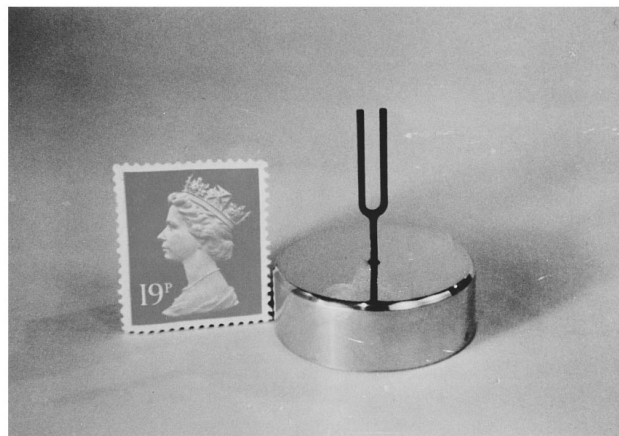
FIG. 2. A photograph of a single-crystal silicon tuning fork which was manufactured using a precision milling machine under computer control. A postage stamp is included for scale.

Miniaturized silicon tuning forks were thus manufactured from thin slices of high-purity float-zone silicon (1 and 2 mm thick) which were cut from an ingot of [110] orientation. The ingot consisted of intrinsic silicon, having n-type (phosphorous) doping and a minimum resistivity of 1000 Ω cm. The thin slices of silicon were rigidly secured between two pieces of glass using a temperature-sensitive bonding wax. The glass/silicon/glass structure was then machined using a precision milling machine under computer control (being careful to terminate the milling before cutting through the deepest glass layer). Following this, the temperature-sensitive bonding wax was simply heated to release the silicon forks. The final step involved a chemical etch of the forks in 10% HF; 90% HNO_3 for 4 min in order to eliminate the 40-μm-thick surface damage layer that results from mechanical machining. Note that this process of manufacture is no more involved than for other forms of resonator, although a computer controlled milling machine is a necessity.

Figure 2 is a photograph of one of the silicon tuning forks. A total of four such silicon forks were manufactured, all with the same crystallographic orientation and the same tine width (see Table I for actual dimensions). Their common crystallographic orientation was established using x-ray diffraction: tine axes were $\langle 110 \rangle$ directions; width directions were $\langle 111 \rangle$; and the plane containing both tine axes was {110}. Thus the Young's modulus of interest (i.e., along the length of the tine) was in the $\langle 110 \rangle$ direction, here symbolized as E_{110}.

B. Tuning forks of hydrided and unhydrided Zr–2.5 wt %Nb

Zirconium alloys are technologically important materials, for they serve as pressure tubes in the CANDU (Canada deuterium uranium) nuclear reactor.[39] Such alloys are used as a result of their low capture cross sections for thermal neutrons and their high strengths. Both Zr–2.5 wt %Nb and Zircaloy-2 have been employed as pressure tubes,[39] but both are susceptible to failure through delayed hydrogen cracking.

TABLE I. Room-temperature tuning fork dimensions and fundamental symmetric mode frequencies.

Material	Tine length[a] L_0 (mm)	Width w_0 (mm)	Thickness (mm)	Fundamental frequency, $f_{1,0}(T_0)$	
				Theory (Hz)	Experiment (Hz)
Silicon	15.0±0.2	0.85±0.05	2.0	5180±440	5410
	15.0±0.2	0.85±0.05	1.0	5180±440	5290
	15.0±0.2	0.85±0.05	1.0	5180±440	5250
	11.3±0.2	0.85±0.05	1.0	9130±860	8570
Zr–2.5%Nb	11.6±0.2	1.00±0.05	1.0	4500±380	4270
(unhydrided)	11.6±0.2	1.00±0.05	1.0	4500±380	4580
Zr–2.5%Nb	11.6±0.2	1.00±0.05	1.0	4500±380	4200
(178 ppm)	11.6±0.2	1.00±0.05	1.0	4500±380	4470

[a]Measurements of L_0 include the distance along the curved portion of the tines.

As this susceptibility can compromise reactor safety, the role of hydrogen in zirconium alloys has been the subject of intense study.[39–47]

Delayed hydrogen cracking occurs in the following way. Zirconium alloys oxidize in the lithiated heavy water which is used as the coolant/moderator in the CANDU reactor. This corrosion reaction also produces hydrogen, ~5% of which diffuses into the tube at the reactor's operating temperature of 300 °C. During 30 years of service the hydrogen concentration in a typical Zr–2.5%Nb pressure tube rises from 10–15 ppm (upon installation) to 50 ppm,[40] although much higher concentrations may result in the presence of stress or temperature gradients.[39,41,42] At low concentrations, the hydrogen remains in solid solution and is not a problem, but upon reaching a critical value the hydrogen begins to precipitate as small solid platelets of zirconium hydride (Zr–1.7 wt %H). These small platelets (of ~10-μm diameter), being brittle, reduce the ultimate tensile strength of the alloy. If the operating temperature of the pressure tube decreases, or if localized stress gradients exist, then an increase in hydride density can result. With time, cracking can occur in regions of hydride formation and pressure tube failures can ensue.[39,43]

Investigations of the effects of hydrogen on the material properties of Zr–2.5%Nb are therefore of considerable interest for reactor safety and reliability. In particular, it was thought that the vibrations of Zr–2.5%Nb tuning forks might provide a convenient means to study the effects of hydrogen on Young's modulus. To this end, four miniaturized tuning forks were manufactured from a small sample of Zr–2.5 wt %Nb pressure tube (~4-mm thick). This sample was prepared by Ontario Hydro (Canada) in such a way as to contain a gradient in hydride concentration along its length, from <15 ppm at one end (the unhydrided end) to 178 ppm±5 ppm at the other. So as to provide a comparison between hydrided and unhydrided material, two tuning forks were machined from each end of the sample such that the hydrogen concentration was spatially constant within any one tuning fork. The machining was again accomplished using the computer-controlled milling machine, but in this case the supporting glass layers were not required. The tine axes were coincident with the axial direction of the pressure tube, while the width (w) directions of the tines lay in the radial (or thickness) direction of the pressure tube. The dimensions

of the Zr–2.5%Nb tuning forks were all the same (see Table I), and were similar to those made of silicon.

In order to model the temperature dependence of Zr–2.5%Nb tuning fork frequencies (see below), some background information regarding the solubility of hydrogen in zirconium alloys will be needed. The required information can be summarized in reference to Fig. 3, which plots typical curves for the terminal solid solubility (TSS) of hydrogen in Zr–2.5%Nb.[39] These TSS curves allow the concentration of dissolved hydrogen to be determined as a function of temperature given the total concentration of hydrogen in the alloy, C_{tot}. Two TSS curves are shown because different responses are known to exist, depending upon whether the hydrides are undergoing dissolution (TSS$_d$) or precipitation (TSS$_p$). In the region to the left of the TSS curves the hydrogen exists in two phases: free (or dissolved) hydrogen in solution, and solid hydrides. In the region to the right of the TSS curves all of the hydrogen exists in the single free-hydrogen phase (i.e., the free-hydrogen concentration equals C_{tot}).

Note that when the material is in the two phase region of Fig. 3, the concentration of free-hydrogen C^f is given by the
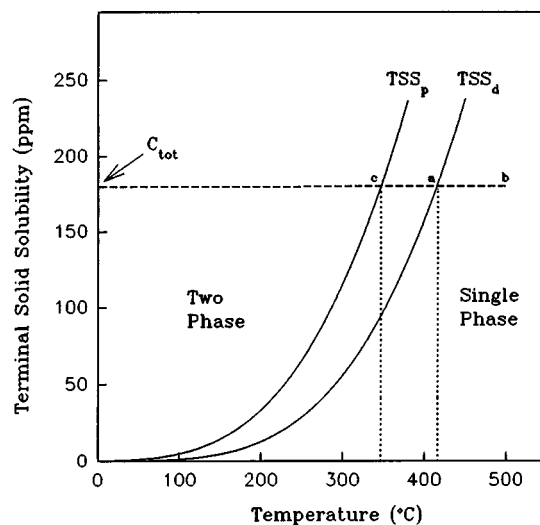
FIG. 3. Typical terminal solid solubility curves of hydrogen in Zr–2.5 wt % Nb.
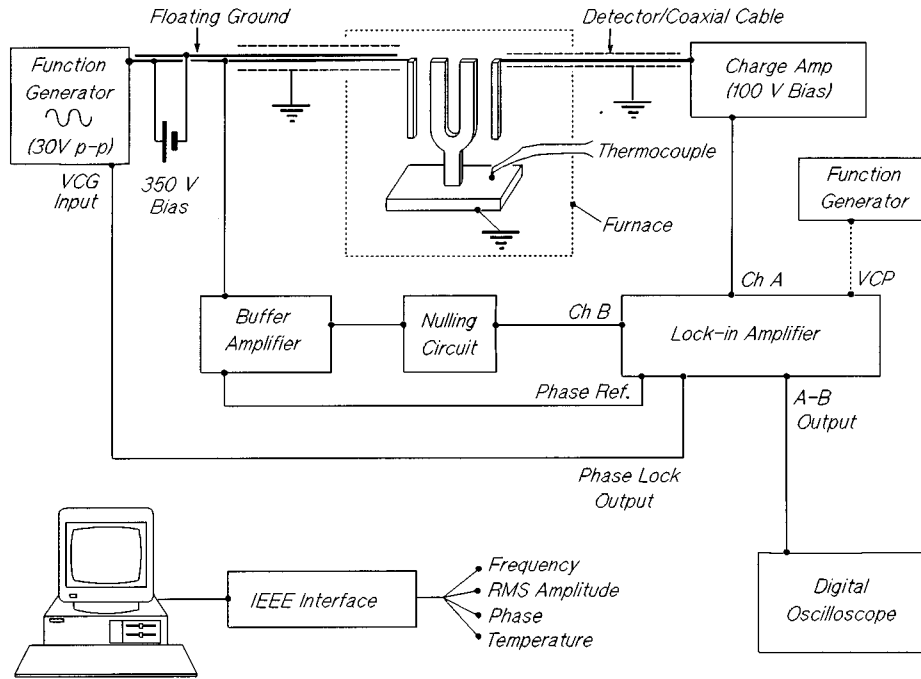
FIG. 4. Schematic diagram of the capacitive system used in experimental studies of continuously driven tuning forks.

explicit functional dependence of the TSS curves as

$$C_i^f[\text{ppm}] = A_i e^{-B_i/T}, \quad i = p, d, \qquad (8)$$

where A_i and B_i are constants, T is the absolute temperature in Kelvin, and $C^f \leq C_{\text{tot}}$. The subscript i is employed to distinguish between the different responses for dissolution and precipitation. The constants A_i and B_i depend upon the particular alloy and its microstructure,[41,45,46] with the curves of Fig. 3 employing average values for Zr–2.5%Nb: $A_p = 41,100$ ppm, $B_p = 3368$ K, $A_d = 60,500$ ppm, and $B_d = 4005$ K.[39]

As an example of how Fig. 3 is used to obtain the free-hydrogen concentration, consider the example of the hydrided Zr–2.5%Nb tuning fork ($C_{\text{tot}} = 178$ ppm) which is cycled in temperature through 500 °C. At room temperature the solubility of hydrogen is negligible and all of the hydrogen exists in the form of precipitated hydrides. As the temperature rises, the solubility increases according to TSS$_d$ and hydrides dissolve to form free hydrogen. When the concentration of free-hydrogen C^f reaches 178 ppm at $\sim$420 °C (point a in the figure) no more hydrides remain and C^f remains constant at 178 ppm for further increases in temperature (line ab). Upon cooling, C^f similarly remains constant until hydrides can begin to precipitate. Precipitation, however, is governed by the TSS$_p$ curve, such that $C^f = 178$ ppm until $\sim$360 °C is reached (point c). Continued cooling then takes C^f back towards zero along TSS$_p$, and at room temperature all of the hydrogen is again in the form of hydrides. This example shows how the free hydrogen concentration follows a hysteresis loop with cycling temperature. Such hysteresis will also be observed in the tuning fork frequency responses below, since the material properties of Zr–2.5%Nb will be shown to depend upon free-hydrogen concentration.

IV. EXPERIMENTAL

The noncontact system used to investigate the vibrations of the miniaturized tuning forks is shown in the schematic diagram of Fig. 4. A continuously driven, capacitive source electrode was used to excite vibrations at the left tine of the tuning fork, while detection was accomplished by a second capacitive electrode at the other tine. The source and receive electrodes consisted of thin metal plates which were placed less than 0.5 mm from the respective tines of the tuning fork. These metal plates were connected to their electronics through specially designed high-temperature coaxial cables. Each coaxial cable consisted of a stainless-steel ground shield surrounding an insulating alumina bar, through which ran the central electrode (see photograph of Fig. 5). The co-
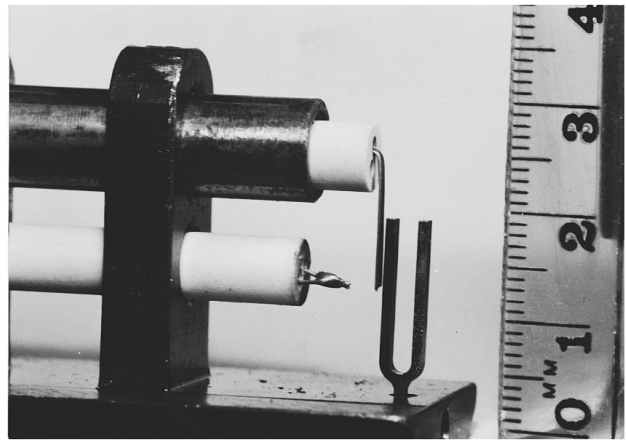
FIG. 5. A photograph of the specially designed high-temperature holder which was used for capacitive detection of tuning fork vibrations. The coaxial cable is the top horizontal tube extending away to the left, and the thermocouple is directly below.
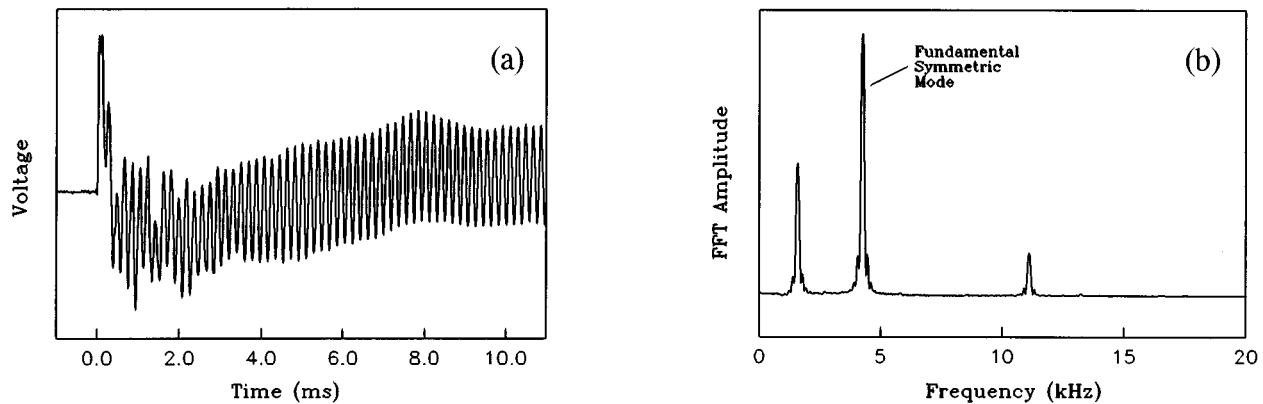
FIG. 6. (a) Typical transient response of a miniaturized tuning fork and (b) typical frequency response. These results were obtained with a Zr–2.5%Nb tuning fork.

axial cable of the source also contained a second, electrically isolated central conductor; this served as a floating ground to reduce ground loops that otherwise occurred when the drive and receive grounds were bolted together.

The drive voltage applied to the source consisted of a sinusoidal voltage from a variable-frequency function generator (Wavetek 191, 30V p-p) in conjunction with a superimposed 350-V bias. This arrangement applied an electrostatic force to the tuning fork at the same frequency as the function generator. Altering the frequency of the function generator then allowed the resonant frequencies of the tuning forks to be excited. (For background on electrostatic transduction see Refs. 18 and 19.)

The detection electrode was electrically connected to a charge sensitive amplifier (Cooknell CA6/C/SU2) which also applied a well-regulated 100-V bias voltage to the detector. Vibration of the grounded tuning fork caused the capacitance of the detector to vary, which, in the presence of the 100-V bias, generated a varying charge. This varying charge was then amplified by the charge-sensitive amplifier and output as a voltage signal. The output voltage from the charge amplifier was sent to channel-A of a phase-sensitive Ithaco 353 lock-in amplifier, whose output in turn was used to form a phase locked loop with the voltage-controlled-gain (VCG) input of the source's function generator. By this method, the steady-state response of the tuning fork could be stabilized under conditions of continuous excitation. In particular, when the phase control of the lock-in amplifier was set to 90°, a resonance frequency of the tuning fork could be tracked as a function of temperature.

One difficulty with a fully capacitive system is electromagnetic coupling between the source and receive electrodes which, in this case, resulted in an output voltage from the charge amplifier at the same phase as the drive voltage. This unwanted signal was removed by using a specially designed buffer amplifier and nulling circuit, in conjunction with differential amplification at the phase-sensitive detector (see Fig. 4).

The temperature of the tuning fork was measured by a Chromel–Alumel (type K) thermocouple, which was placed below and parallel to the detector's high-temperature coaxial cable. The end of the thermocouple was inserted into a small vertical hole that was drilled into the stainless-steel base near the tuning fork (<5 mm away), and carbon filings were packed into the intervening space to provide improved thermal contact. The output voltage from the thermocouple was recorded by a digital voltmeter and converted to temperature using the standard 8th-order, power-series polynomial for a K-type thermocouple. An Omega electronic ice-point reference provided the 0 °C reference temperature, and temperature measurements were considered to be accurate to ±1 °C.

All data acquisition in the system was automated using a PC computer and an IEEE interface. In particular, three Keithley-175 digital voltmeters, a frequency meter (Philips PM6671), and a digital oscilloscope (Tektronix 2430 A) were interfaced to the computer. These instruments allowed automated measurements of the tuning fork's amplitude and phase response, as well as its frequency and temperature. Note that the high-temperature components of the system were positioned at the midpoint of a Wild–Barfield cylindrical furnace, the temperature of which was well-regulated by a programmable temperature controller (Omega CN-20100) for provision of predetermined temperature-time profiles.

V. RESULTS

A. Initial measurements at room temperature

An initial experiment investigated the transient responses of the miniaturized tuning forks, using a pulsed laser for excitation and the capacitance detector. The radiation from a Q-switched ruby laser (30-ns pulsewidth), was directed at one of the tines of a tuning fork so as to create a wideband thermoelastic source.[48] The transient source excited the tuning fork into a superposition of its normal modes of vibration, which decayed in a damped-free manner. A typical transient response of a Zr–2.5%Nb tuning fork is shown in Fig. 6(a). The radiation from the pulsed laser arrives at $t=0$ and excites the tuning fork into vibration. Initially, the vibrations produce a complex temporal response in the detector but, with time, the complexity dies away and the response approaches a high-Q sinusoidal vibration.

The initial vibrations consist of a superposition of three vibrational modes, as shown in the frequency response of
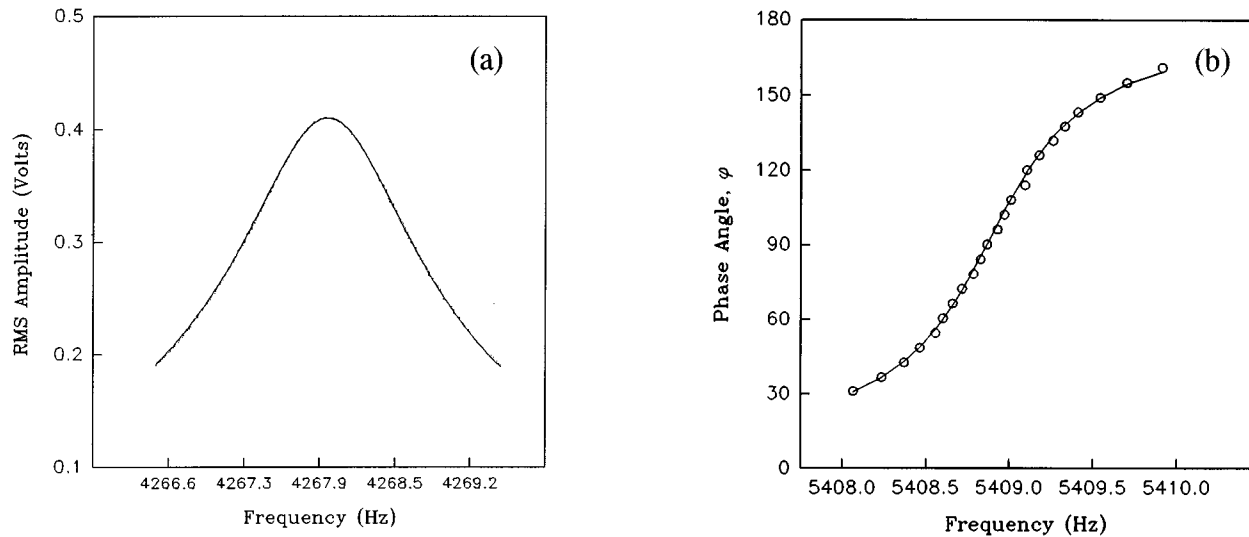
FIG. 7. Typical amplitude and phase responses at room temperature for continuously driven tuning forks near their fundamental symmetric mode frequencies. In (a) the amplitude response of a Zr–2.5%Nb fork is given, while in (b) the phase response of a silicon tuning fork is shown.

Fig. 6(b). This spectrum was obtained by a fast Fourier transform of the transient temporal response using a time window that included the initial, more complex temporal region. The lowest frequency peak is the fundamental cantilever-beamlike mode, in which the whole tuning fork sways back and forth within the plane of the tines. The second (or middle) peak is the fundamental symmetric mode of vibration $f_{1,0}$. The upper mode is probably either the fundamental out-of-plane cantilever-beamlike mode, or the fundamental torsional-like mode (depending upon the particular tuning fork employed). Both the upper and lower frequency modes decay rapidly with time, and contribute only to the initial more complex portion of the temporal response. The fundamental symmetric mode, on the other hand, remains to create the high-Q sinusoidal vibration. This symmetric mode was selected for use in the continuous excitation system to study tuning fork vibrations as a function of temperature.

The spectra that resulted from laser-excited transient waveforms, for all of the tuning forks studied, were used to obtain room-temperature values of fundamental symmetric mode frequencies for comparison with theoretical expectations. The results are summarized in Table I. The theoretical values were calculated using Eq. (1), given the measured dimensions of the tuning forks and known data for the material properties: $E_{110}(T_0) = 1.6776 \times 10^{11}$ Pa (Ref. 12) and $\rho_0 = 2.33$ g/cm^3 (Ref. 34) for silicon and $E(T_0) = 91$ GPa and $\rho_0 = 6.48$ g/cm^3 (Ref. 44) for Zr–2.5%Nb. Both sets of measured results can be seen to agree with theory (within the error limits on the theoretical values). Note that the frequencies are much greater than the calculated thermal-wave frequency within the tine width.[2] This means that estimates of elastic moduli from measurements of symmetric mode frequencies will yield the adiabatic values.

The fundamental symmetric mode vibrations were also investigated at room temperature using the fully capacitive phase-sensitive detection scheme of Fig. 4. Typical amplitude and phase responses at room temperature are shown in

Fig. 7(a) and (b) for the fundamental symmetric modes. Figure 7(a) plots raw data for the amplitude response of a Zr–2.5%Nb tuning fork, along with a best-fit line for the classic amplitude response of a resistively damped resonance, $A(f) = A_0 f_0 / f [(f_0/f - f/f_0)^2 + 1/Q^2]^{-1/2}$, where the free parameters-of-fit are amplitude A_0, quality factor Q, and resonance frequency f_0. As the best-fit line cannot be visually discerned from the data, it is concluded that the tuning fork is well-approximated by a linear, damped harmonic oscillator with damping proportional to velocity. The best-fit parameters of interest were $f_{1,0}(T_0) = 4267.963 \pm 0.001$ and $Q(T_0) = 2805.9 \pm 1.5$, with a full width at half maximum of ~ 1.5 Hz. These results demonstrate the high Q's of the tuning forks, and the high accuracy that is available with the phase-sensitive system (i.e., less than 1 part in 10^6 for measurements of resonance frequency).

Similar results were obtained from an analysis of the phase response of Fig. 7(b), which plots the raw data for a silicon tuning fork ($\bigcirc$), along with a best-fit line for the classic phase response through resonance, $\phi(f) = \arctan[Q(f_0/f - f/f_0)]^{-1}$. Again, good agreement is observed, with the fit-parameters given as $f_{1,0} = 5408.882 \pm 0.003$ Hz and $Q = 6195 \pm 111$. Note that for both silicon and Zr–2.5%Nb tuning forks, the Q's of the resonances were so high that a phase-locked loop was an absolute necessity in order to observe the resonances under continuous excitation.

B. Responses of silicon tuning forks at high temperatures

In order to study the fundamental symmetric mode frequencies as a function of temperature, the tuning forks were phase locked at 90° relative to the phase of the driving voltage (i.e., at the point of maximum vibration amplitude). In this way the resonance frequency was simply tracked by the capacitance system as a function of temperature. The nor-
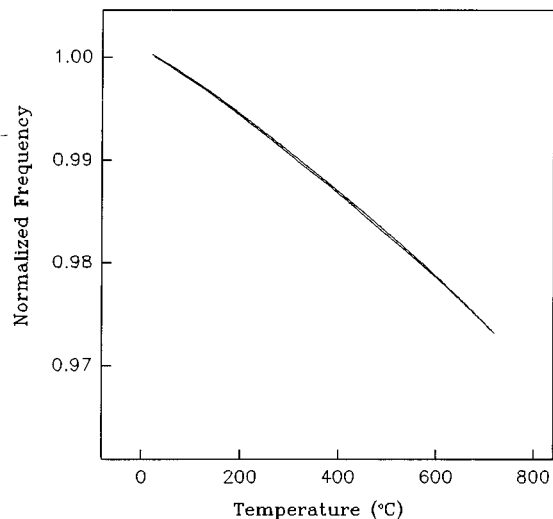
FIG. 8. Typical temperature dependence of the fundamental symmetric mode frequency for a silicon tuning fork, as measured by the phase-sensitive system.

FIG. 9. Temperature dependence of Young's modulus e_{110} for silicon, as obtained from the tuning fork data of Fig. 8 and from the converted data of Burenkov and Nikanorov.[12]

malized results for silicon, for a single temperature cycle ($20° \rightarrow 700° \rightarrow 20\,°C$), are shown in Fig. 8. This data, which is typical of all the silicon forks studied, was acquired at $0.5\,°C$ intervals in order to produce the continuous curve shown. The frequency of the fundamental is seen to fall by $\sim 2.5\%$ over $700\,°C$, where the smooth nature of the curve is readily apparent. Such smooth and repeatable results suggest that the silicon tuning fork could serve as a secondary thermometer by calibrating its frequency response against a primary standard. The slight separation of the increasing and decreasing temperature data at the curve midpoints is a result of the tuning fork temperature lagging behind the thermocouple response by fractions of a degree. This difference does not appear at the end points of the curves because the rate of temperature change was reduced in these regions. The fact that the system has detected this difference at all demonstrates its ability to resolve fine structures in temperature responses. In fact, such resolution would appear to be limited by the accuracy of the frequency measurements, which is of the order of the line thickness in the figure.

The frequency data of Fig. 8 was used to determine the normalized temperature dependence of Young's modulus for silicon, $e_{110}(T)$, using Eq. (7). This conversion was possible given $\alpha_v(T) = 3\alpha_l(T)$ for single crystals with cubic symmetry,[33,35–37] and given known data for $\alpha_l(T)$.[35,38] The results are included in Fig. 9 (●) where the accuracy was limited not by the thermal expansion data, but by the use of the thin-tine approximation ($<0.05\%$).

Also included in Fig. 9, for comparison, is the expected temperature dependence of Young's modulus of silicon, which was calculated in the following way. As silicon is a cubic crystal with a diamond structure, its elasticity tensor is defined by three independent elastic compliances s_{11}, s_{12}, and s_{44}. Using a prescription for evaluating Young's modulus along any direction of a cubic crystal in terms of the compliance matrix (Ref. 34, pp. 213–221), Young's modulus along the tine length can be expressed as,
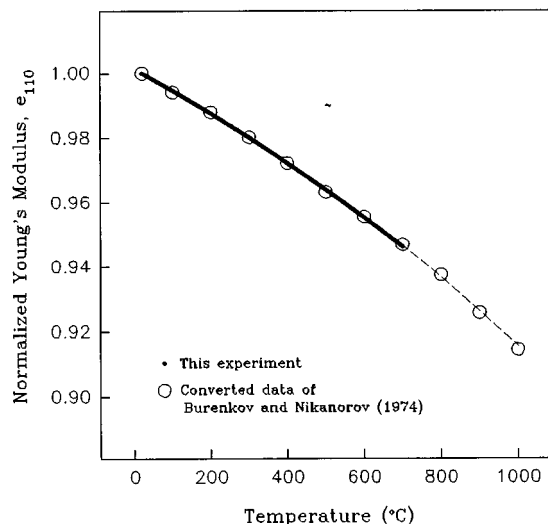
$$E_{110} = \frac{2}{(s_{11} + s_{12} + \frac{1}{2}s_{44})}. \tag{9}$$

This equation allowed E_{110} to be determined at any temperature using the data of Burenkov and Nikanorov[12] for the temperature variation of elastic compliances $s_{ij}(T)$. With $E_{110}(T)$ determined in this way, $e_{110}(T)$ was simply obtained through division by the room-temperature Young's modulus, $E_{110}(T_0) = 1.6776 \times 10^{11}$ N/m^2. Note that the comparison results have a quoted accuracy of 0.4%[12] (i.e., an order of magnitude less accurate than the tuning fork data). The agreement between the measured and expected results is excellent, with the data of this study following closely the best-fit curve for the data of Burenkov and Nikanorov (dashed line).

Three important conclusions can be drawn from the agreement in Fig. 9. First, the agreement provides independent verification of the results of Burenkov and Nikanorov. Second, the agreement confirms the temperature calibration of the tuning fork system and the thin-tine approximation assumed in the theory; this allows accurate studies of other materials to proceed with confidence. Finally, as the scatter in the comparison data is greater than the tuning fork results, it can be concluded that the noncontact tuning fork system has performed more accurately than that employed by Burenkov and Nikanorov. This is not surprising given the very high Q's of tuning fork resonators and their inherent isolation from supporting clamps (i.e., as compared with the longitudinal bulk-wave resonances of cylindrical bars used by Burenkov and Nikanorov).

Note that these three conclusions would have been difficult to support if not for the unparalleled purity and crystalline perfection with which silicon is produced. After all, differences in material properties did not have to be considered, and results from the literature could be accurately employed for the calibration and comparison of the two systems. This strongly suggests that single crystal silicon should
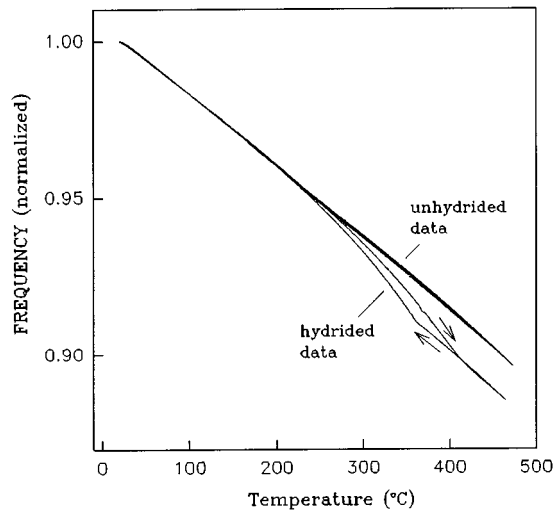
FIG. 10. Temperature dependencies of fundamental symmetric mode frequencies for Zr–2.5%Nb tuning forks, as measured by the phase-sensitive system.

FIG. 11. Temperature dependencies of normalized Young's moduli for hydrided and unhydrided Zr–2.5%Nb tuning forks, as obtained from the data of Fig. 10.

be employed as a standard material for the calibration of high-temperature acoustical measurement systems.

C. Responses of Zr−2.5%Nb tuning forks at high temperatures

Measurements of fundamental symmetric mode frequencies for hydrided and unhydrided Zr–2.5%Nb tuning forks are shown in Fig. 10 for a single-temperature cycle ($20° \rightarrow 470° \rightarrow 20$ C over a period of 8 h). These normalized results were reproducible with both sets of tuning forks. The unhydrided tuning fork frequencies are seen to exhibit an approximately linear single-valued response, regardless of heating or cooling, whereas the hydrided tuning forks trace out a hysteresis loop indicated by the arrows (top curve for heating, bottom curve for cooling). The hysteresis loop of the hydrided data was expected (as mentioned earlier) due to effects of dissolved hydrogen on the material properties of Zr–2.5%Nb.

The frequency data of Fig. 10 was converted to normalized Young's moduli using Eq. (7), and the results are presented in Fig. 11. The conversion assumed that $\alpha_v(T) = 3\alpha_l(T)$, and that the thermal expansion coefficients were isotropic and temperature independent [$\alpha_l^{uh} = 6.6 \times 10^{-6}$ °C^{-1} (Ref. 44) for unhydrided material and $\alpha_l^h = 6.1 \times 10^{-6}$ °C^{-1} for hydrided material (Ref. 47, Fig. 5)]. Note that the assumption of isotropic, temperature-independent, thermal-expansion coefficients is not strictly valid, since the coefficients are known to exhibit a mild anisotropy of $\sim 5\%$ and a small temperature dependence, whereas hydrided Zr–2.5%Nb also exhibits an additional minor hysteresis loop.[44,45,47] These small nonlinearities, however, can be ignored for the present purpose without significant penalty ($\leqslant 0.2\%$ error in Fig. 11).

The curves of Fig. 11 can be properly interpreted by assuming that Young's modulus of Zr–2.5%Nb decreases as the free hydrogen concentration in the alloy increases. To see this, it is useful to follow the predictions of this assumption
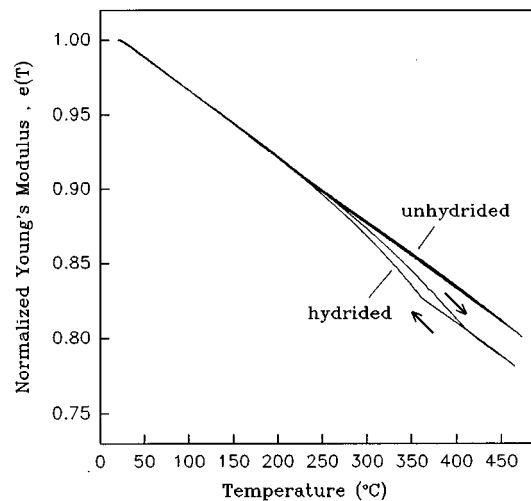
for a single temperature cycle while comparing the predictions to Fig. 11. When the temperature is low and rising, the hydrided modulus $e^h(T)$ decreases with temperature in the same manner as the unhydrided modulus $e^{uh}(T)$. This occurs because the solubility of hydrogen is low (see Fig. 3) and the free-hydrogen concentration C^f is not sufficient to alter Young's modulus. As the temperature continues to rise, the free-hydrogen concentration of the hydrided fork increases exponentially according to the terminal solid-solubility curve for dissociation (TSS$_d$), and this decreases $e^h(T)$ accordingly (see $\searrow$ in Fig. 11). When the free-hydrogen concentration reaches a maximum of C_{tot} at ~ 420 °C, all of the hydrides are dissolved and the source for free hydrogen is exhausted; this results in the abrupt change of slope in the hydrided curve. At temperatures beyond the abrupt transition, the free-hydrogen concentration remains constant and the hydrided modulus decreases along a line that is approximately parallel to the unhydrided modulus. Upon cooling, the process simply reverses, except now the abrupt change in slope occurs at a lower temperature of ~ 360 °C which is consistent with hydride precipitation (i.e., determined from the TSS$_p$ curve with $C_p^f = C_{tot}$). Continued cooling then brings C^f back toward zero according to TSS$_p$ ($\diagdown$), and the hydrided modulus returns to its original value. (Note that the unhydrided fork experiences a simple linear softening with temperature since it does not contain appreciable hydrogen.)

D. Quantitative analysis of Zr−2.5%Nb tuning fork responses

The above interpretation of Fig. 11 can be made quantitative by assuming that Young's modulus of Zr–2.5%Nb decreases in proportion to both free hydrogen concentration and temperature.[44] Thus with Young's modulus of hydrided Zr–2.5%Nb symbolized as $E^h(T)$, the normalized temperature dependence $e^h(T)$ can be expressed as
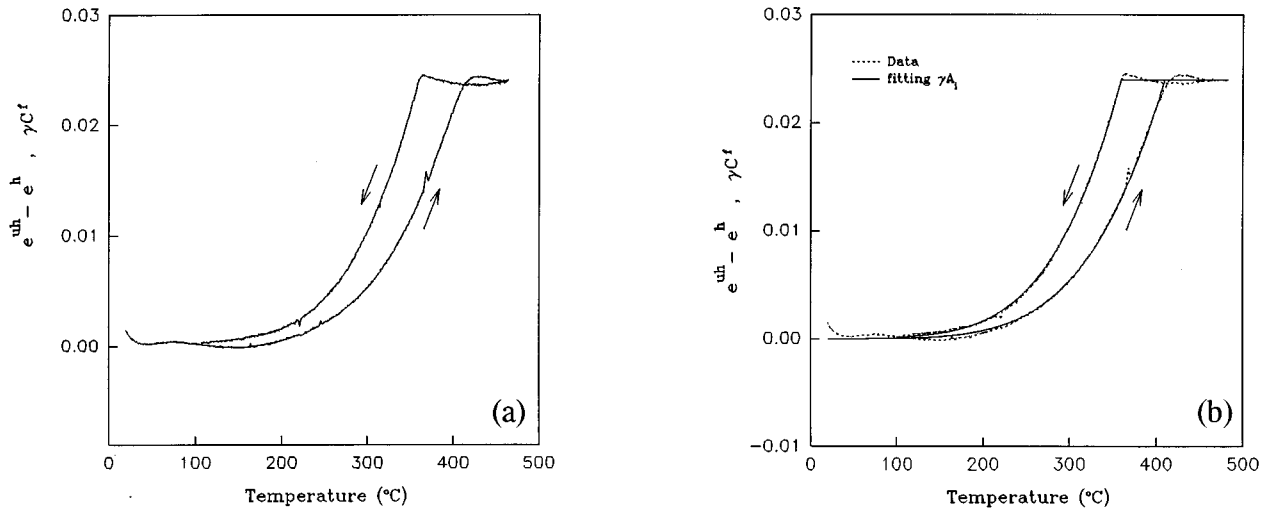
FIG. 12. The difference between normalized Young's moduli for hydrided and unhydrided Zr–2.5%Nb tuning forks. The converted data is shown in (a) and is fit to the equation for free-hydrogen concentration in (b).

$$e^h(T) = \frac{E^h(T)}{E_0^h} = [1 - \beta(T - T_0) - \gamma C_i^f], \qquad (10)$$

where β is the proportionality constant that accounts for the normal linear temperature variation (in the absence of hydrogen), and γ is the proportionality constant that accounts for the effects of hydrogen. This equation shows that for hydrided material under conditions of cycling temperature, Young's modulus (and thus resonance tuning fork frequencies) follow hysteresis loops along with C^f. As for the normalized Young's modulus of unhydrided Zr–2.5%Nb, $e^{uh}(T)$, it can be obtained from Eq. (10) by ignoring the third term in the brackets, such that $e^{uh}(T) = 1 - \beta(T - T_0)$.

The proportionality constant of the temperature term β was easily obtained by a linear regression of the unhydrided data of Fig. 11: $\beta = (4.410 \pm 0.009) \times 10^{-4} \, {}^\circ\text{C}^{-1}$. Multiplying this value of β by the room-temperature Young's modulus ($E_0 = 91$ GPa) shows that the absolute value of Young's modulus for unhydrided Zr–2.5%Nb, $E^{uh}(T)$, softened with temperature during experiment at a rate of 40.13 MPa/°C.

As for the other proportionality constant in Eq. (10), γ, it is somewhat more difficult to obtain and requires the following circuitous analysis. First, note that the difference between the normalized Young's moduli for hydrided and unhydrided tuning forks is

$$e^{uh}(T) - e^h(T) = \gamma C_i^f, \qquad (11)$$

which shows the difference to be proportional to free hydrogen concentration. This difference was obtained from the experimental data of Fig. 11 and is plotted in Fig. 12(a), where the result is reminiscent of the free-hydrogen hysteresis loop, as described earlier in reference to Fig. 3.

Figure 12(a) can indeed be accurately interpreted as the free-hydrogen concentration in the hydrided tuning fork during the course of the temperature cycle (apart from a scaling by the coefficient γ). To prove this, the data of Fig. 12(a) was fit to the expected exponential form of TSS_i curves,

using Eq. (8) and the following best-fit parameters: $B_d = 5323 \pm 53$ K, $B_p = 5115 \pm 51$ K, $\gamma A_d = 58.2 \pm 0.6$ and $\gamma A_p = 77.6 \pm 0.8$. The results are plotted in Fig. 12(b), where the agreement is obviously excellent. Recall that the right curve indicates the increase in free-hydrogen concentration C^f as the tuning fork's temperature rises (i.e., TSS_d) whereas the left curve indicates the decrease in C^f that results as the hydrides reprecipitate upon cooling (i.e., TSS_p). The plateau that occurs above $\sim 420\,^\circ\text{C}$ for TSS_d, and above $\sim 360\,^\circ\text{C}$ for TSS_p, is the region in which all of the hydrides are dissolved, and the height of this plateau is related to the total hydrogen concentration in the material. When the height of the plateau region in Fig. 12(a) is multiplied by E_0, a value of 2.1 GPa results. This is the value of the constant offset that occurs in the hydrided Young's modulus beyond the transition region of Fig. 11 (as measured relative to the unhydrided modulus). Such excellent agreement between theory and experiment confirms that the difference of the moduli $e^{uh}(T) - e^h(T)$ is indeed a measure of free-hydrogen concentration, and lends strong support to the model that was used to predict Young's modulus of Zr–2.5%Nb in the presence of hydrides.

With the curves of Fig. 12 now proven to be proportional to TSS_i curves, the final step can be taken toward obtaining the coefficient γ. As the total hydrogen concentration of the hydrided tuning fork is known ($C_{\text{tot}} = 178$ ppm ± 5 ppm) and as the plateau region of Fig. 12(a) occurred at $\gamma C_{\text{tot}} = 0.0239 \pm 0.0003$, the value of γ is simply determined by division to be $\gamma = (1.34 \pm 0.05) \times 10^{-4}$ ppm^{-1}. Given this value, the A_i coefficients of Eq. (8) can now be evaluated using the above results for γA_i [$A_d = (4.3 \pm 0.2) \times 10^5$ ppm and $A_p = (5.8 \pm 0.3) \times 10^5$ ppm], such that unscaled TSS_i curves for hydrided Zr–2.5%Nb can be predicted from the experimental data. These curves are plotted in Fig. 13 (solid lines), along with a second set of typical TSS_i curves for Zr–2.5%Nb as obtained by Byrne and Leger[39] (dotted lines). The obvious differences between the
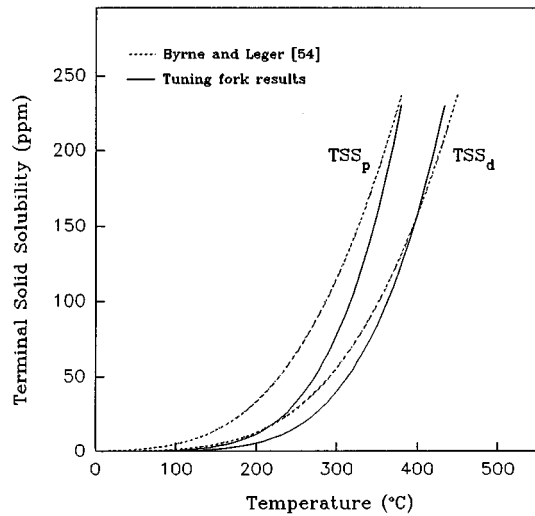
FIG. 13. TSS_i curves as obtained from a numerical analysis of the Zr-2.5% Nb tuning fork measurements. The curves of Byrne and Leger[39] are also shown for comparison.

two sets of TSS_i curves are attributed to differences in samples, which is justified since TSS_i curves are well-known to be sensitive to the particular microstructure and alloy composition[41,45,46] and to the heating/cooling rates employed during investigations.[47] As effects upon microstructure may have resulted when the small tuning forks were milled, and as heating and cooling rates have not been variables of study to date, further work is necessary to quantify the observed differences in TSS_i curves. Completing such work would no doubt be useful, for it may be that it is the past methods of measuring TSS_i curves that are in error. In any case, the fact that the capacitive tuning fork system was able to determine the entire temperature dependence of TSS_i curves from the frequency responses of two tuning forks is of considerable interest. After all, other methods that seek to measure the same (such as dilatometry methods) must employ a host of

samples that cover a wide range of hydrogen concentrations (see, for example, Ref. 45).

The quantitative analysis of Zr–2.5%Nb tuning fork responses was taken one final step in order to generate the theoretical frequencies of the tuning forks, so as to compare with the initial raw data. This was possible since all of the parameters in Eq. (10) have been accurately specified through the above analysis of experimental data. The result is shown in Fig. 14(a), where the agreement, once again, is excellent. As the experimental and theoretical curves cannot be visually distinguished in this figure, the relative percentage error has been included in Fig. 14(b), where the results are separated into heating and cooling curves. Apart from a 0.1% difference upon start-up of the system at $T=20°$ (which was due to a brief heating rate that was too rapid for the temperatures of the thermocouple and tuning fork to be in-sync.), and apart from a number of observed high-frequency transient spikes (due to the lock-in amplifier experiencing brief destabilizations of its phase-locked loop), the percentage difference remains below 0.05%. This indicates the well-characterized response of the experimental system and the high accuracy with which Zr–2.5%Nb Young's moduli have been modeled by a simple theory.

VI. CONCLUSION

The use of noncontact capacitive transducers has allowed resonant vibrations to be investigated, without the damping or stresses that would be introduced by contacting transducers. The use of miniaturized tuning forks in their fundamental symmetric modes of vibration provided resonators with natural nodes that were isolated from their supporting clamps. This eliminated irreversible frequency and drift effects that often occur with other forms of resonator. The fact that the tuning forks had such high Q's resulted in measurements that were highly sensitive to changes in material properties and temperature, and allowed accurate quantitative analysis of the effects of hydrogen on Young's modulus
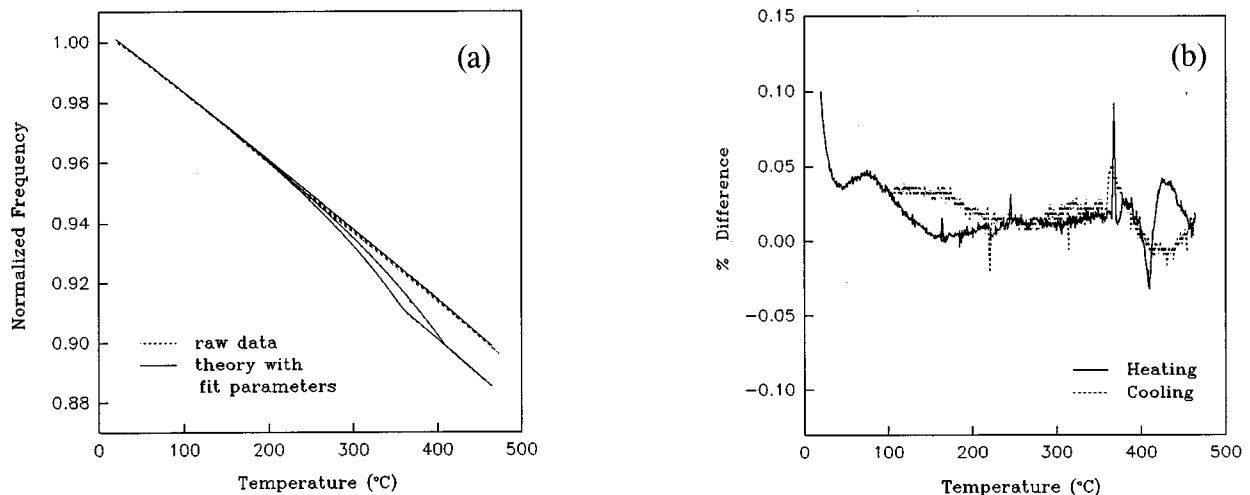
FIG. 14. (a) Theoretical frequency results obtained from the experimentally obtained fit parameters, as compared to the raw data; (b) the relative percentage difference between the theoretical and experimental hydrided data in (a).

in Zr–2.5%Nb. The accuracy and reliability of the measurements were confirmed by the silicon tuning fork results which provided an absolute calibration of the acoustical system through a comparison with known data for the elastic moduli of silicon.

One point to note is the tuning fork system was only capable of measuring Young's modulus in the direction of the tine length. This is somewhat limiting, as a number of different tuning forks would need to be manufactured in order to determine all of the extensional elastic moduli of an anisotropic material, whereas shear moduli cannot be determined at all. However, as was seen in the case of single-crystal silicon, Young's modulus in an anisotropic solid will generally depend upon all of the elastic constants of the material (including the shear moduli). This means that a single tuning fork of suitable orientation can still be used to verify other measurement techniques that are capable of measuring the entire elastic tensor of an anisotropic solid as a function of temperature. One such method is the rectangular parallel-epiped resonance (or RPR) method, as discussed in the Introduction, whose results are presently in need of independent confirmation. The developed tuning fork system could therefore provide the needed calibration. (Other avenues for future research and sensor developments have been discussed by the authors elsewhere.[18])

ACKNOWLEDGMENTS

[1] T.-S. Kê, "Experimental evidence of the viscous behaviour of grain boundaries in metals," Phys. Rev. **71**, 533–546 (1947).

[2] M. Barmatz and B. Golding, "Thermoelastic relaxation near the Curie point of EuO," Phys. Rev. B **9**, 3064–3073 (1974).

[3] M. Barmatz and H. S. Chen, "Young's modulus and internal friction in metallic glass alloys from 1.5° to 300 °K," Phys. Rev. B **9**, 4073–4083 (1974).

[4] T. Tiedje, R. R. Haering, and W. N. Hardy, "The application of capacitive transducers to sound velocity measurements in TTF-TCNQ," J. Acoust. Soc. Am. **65**, 1171–1181 (1979).

[5] C. A. Wert, "Internal friction in solids," J. Appl. Phys. **60**, 1888–1895 (1986).

[6] C. Zener, *Elasticity and Anelasticity of Metals* (University of Chicago, Chicago, 1948).

[7] J. B. Wachtman, Jr., W. E. Tefft, D. G. Lam, Jr., and R. P. Stinchfield, "Elastic constants of synthetic single crystal corundum at room temperature," J. Res. Natl. Bur. Stand. Sec. A **64A**, 213–228 (1960).

[8] S. Spinner and W. E. Tefft, "A method for determining mechanical resonance frequencies and for calculating elastic moduli from these frequencies," Proc. ASTM **64**, 1221–1238 (1961).

[9] A. Migliori, J. L. Sarrao, W. M. Visscher, T. M. Bell, M. Lei, Z. Fisk, and R. G. Leisure, "Resonant ultrasound spectroscopic techniques for measurement of the elastic moduli of solids," Physica B **183**, 1–24 (1993).

[10] A. Migliori, W. M. Visscher, S. Wong, S. E. Brown, I. Tanaka, H. Kojima, and P. B. Allen, "Complete elastic constants and giant softening of c_{66} in superconducting La$_{1.86}$Sr$_{0.14}$CuO$_4$," Phys. Rev. Lett. **64**, 2458–2461 (1990).

[11] A. I. Drozzhin, S. A. Antipov, and A. M. Belikov, "High-temperature damping and microplasticity of silicon," Sov. Phys. Solid State **24**, 693–694 (1982).

[12] Yu. A. Burenkov and S. P. Nikanorov, "Temperature dependence of elastic constants of silicon," Sov. Phys. Solid State **16**, 963–964 (1974).

[13] S. Spinner, "Temperature dependence of elastic constants of some cermet specimens," J. Res. Natl. Bur. Stand. Sec. C Eng. Instrum. **65**, 89–96 (1961).

[14] Y. Sumino, O. L. Anderson, and I. Suzuki, "Temperature coefficients of elastic constants of single crystal MgO between 80 and 1,300 K," Phys. Chem. Miner. **9**, 38–47 (1983).

[15] O. L. Anderson and K. Zou, "Thermodynamic functions and properties of MgO at high compression and high temperature," J. Phys. Chem. Ref. Data **19**, 69–83 (1990).

[16] I. Suzuki, O. L. Anderson, and Y. Sumino, "Elastic properties of a single-crystal forsterite Mg$_2$SiO$_4$, up to 1,200 K," Phys. Chem. Miner. **10**, 38–46 (1983).

[17] T. Goto and O. L. Anderson, "Apparatus for measuring elastic constants of single crystals by a resonance technique up to 1825 K," Rev. Sci. Instrum. **59**, 1405–1408 (1988).

[18] D. W. Schindel, "Progress in non-contact acoustic methods," Ph.D. thesis, Department of Physics, Queen's University, Canada, 1995.

[19] D. W. Schindel and D. A. Hutchins, "The capacitance transducer as a standard source of ultrasound in solids," J. Acoust. Soc. Am. **97**, 1650–1659 (1995).

[20] D. W. Schindel, D. A. Hutchins, and S. T. Smith, "A study of materials at high temperature using miniaturized resonant tuning forks," The 124th Meeting of the Acoustical Society of America, New Orleans, LA, 31 Oct.–4 Nov. 1992 [J. Acoust. Soc. Am. **92**, 2314(A) (1992)].

[21] D. A. Hutchins, C. Edwards, D. W. Schindel, M. D. C. Moles, and H. D. Mair, "Laser-based high-temperature ultrasonic studies of hydrided zirconium alloys," in *IEEE 1991 Ultrasonics Symposium Proceedings*, December 1991 (IEEE, New York, 1991), Vol. 2, pp. 907–910.

[22] A. Wood, *The Physics of Music* (Wiley, New York, 1975), pp. 148–149.

[23] J. Tyndall, *The Science of Sound* (Philosophical Library, New York, 1964).

[24] Lord Rayleigh, *The Theory of Sound* (Dover, New York, 1945), 2nd ed.

[25] A. B. Wood, *A Textbook of Sound* (Bell, London, 1955), pp. 122–137.

[26] Y. Tomikawa, S. Oyama, and M. Konno, "A quartz crystal tuning fork with modified basewidth for a high quality factor: Finite element analysis and experiments," IEEE Trans. Sonics Ultrason. **SU-29**, 217–223 (1982).

[27] E. Momosaki and T. Kaneko, "A study of quartz tuning fork resonators in the overtone mode with unwanted temperature characteristics," Jpn. J. Appl. Phys. **33**(5B), Part 1, 2994–2997 (1994).

[28] A. A. Fathimani and J. F. W. Bell, "A new resonant thermometer for nuclear reactor applications," J. Phys. E **11**, 588–596 (1978).

[29] J. Söderkvist, "Piezoelectric beams and vibrating angular rate sensors," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 271–280 (1991).

[30] P. L. Kebabian, S. Kallelis, D. D. Nelson, Jr., and A. Freedman, "UHV-compatible electrostatically driven tuning fork chopper," Rev. Sci. Instrum. **64**, 346–348 (1993).

[31] J. Backus, *The Acoustical Foundations of Music* (Norton, New York, 1969), pp. 65–71.

[32] T. D. Rossing, D. A. Russel, and D. E. Brown, "On the acoustics of tuning forks," Am. J. Phys. **60**, 620–626 (1992).

[33] W. D. Callister, Jr., *Materials Science and Engineering* (Wiley, New York, 1985).

[34] W. R. Runyan, *Silicon Semiconductor Technology* (McGraw-Hill, New York, 1965).

[35] H. F. Wolf, *Silicon Semiconductor Data* (Pergamon, London, 1969).

[36] Y. S. Touloukian, R. K. Kirby, R. E. Taylor, and P. D. Desai, *Thermophysical Properties of Matter: Thermal Expansion, Metallic Elements and Alloys* (Plenum, New York, 1977), Vol. 12.

[37] *American Institute of Physics Handbook* (Colonial Press, New York, 1972), pp. 4:119–4:128.

[38] Y. S. Touloukian, *Thermophysical Properties of High Temperature Solid Materials, Volume 1: Elements* (MacMillan, New York, 1967), pp. 878–893.

[39] T. J. Byrne and M. Leger, "Hydride blister growth modelling—Progress report No. 2," Ontario Hydro Internal Report No. 85-29-H, 1985.

[40] B. A. Cheadle, C. E. Coleman, and H. Licht, "CANDU-PHW pressure

tubes: their manufacture, inspection, and properties,'' Nucl. Technol. **57**, 413–425 (1982).

[41] M. P. Puls, ''On the consequences of hydrogen supersaturation effects in Zr alloys to hydrogen ingress and delayed hydride cracking,'' J. Nucl. Mater. **165**, 128–141 (1989).

[42] R. L. Eadie, D. Mok, D. Scarth, and M. Leger, ''The hydrostatic stress field around the crack tip in zirconium 2.5% niobium and implications for delayed hydride cracking,'' Scr. Metall. Mater. **25**, 497–502 (1991).

[43] T. Laursen, M. Leger, X-P. Ma, J. D. MacArthur, G. R. Palmer, and J. L. Whitton, ''The measurement of the deuterium concentration distributions in deuteride blisters on zirconium-alloy pressure tube material,'' J. Nucl. Mater. **165**, 156–163 (1989).

[44] D. Mair (private communication).

[45] G. F. Slattery, ''The terminal solubility of hydrogen in zirconium alloys between 30 and 400 °C,'' J. Inst. Met. **95**, 43–47 (1967).

[46] M. P. Puls, ''The effects of misfit and external stresses on terminal solid solubility in hydride-forming metals,'' Acta Metall. **29**, 1961–1968 (1981).

[47] F. D. Manchester and G. C. Weatherly, ''The solid solubility of hydrogen in Zr–2.5 wt % Nb pressure tube alloys,'' Ontario Hydro Final Contract Report, P.O. No. 18-39318-11 and No. 19-0222-11, 1980, pp. 1–35.

[48] D. A. Hutchins, ''Ultrasonic generation by pulsed lasers,'' in *Physical Acoustics*, edited by R. N. Thurston and A. D. Pierce (Academic, New York, 1988), Vol. XVIII, Chap. 2, pp. 21–123.

An integral equation formulation for boundary element analysis of propagation in viscothermal fluids

C. Karra and M. Ben Tahar

LG2mS, UPRES A CNRS No. 6066, Université de Technologie de Compiègne, BP 29.529, 60205 Compiègne Cedex, France

A new variational formulation using integral equations for the resolution of a propagation problem in a viscous and thermally conducting fluid is investigated. Numerical techniques are applied to a geometry of revolution. Results are compared to an analytic solution for a cylindrical cavity filled with air. These results illustrate the effects of viscosity and thermal conductivity of air in the case of micro cavities. © *1997 Acoustical Society of America.* [S0001-4966(97)03908-8]

PACS numbers: 43.20.Bi, 43.35.Ud [JEG]

LIST OF SYMBOLS

p	pressure variation
p_a	acoustic pressure
p_h	entropic pressure
T	temperature variation
T_a	acoustic temperature
T_h	entropic temperature
$\mathbf{v}$	particle velocity
$\mathbf{v}_e$	irrotational velocity ($\mathrm{curl}\,\mathbf{v}_e = 0$)
$\mathbf{v}_e^a$	acoustic irrotational velocity
$\mathbf{v}_e^h$	entropic irrotational velocity
$\mathbf{v}_t$	rotational velocity ($\mathrm{div}\,\mathbf{v}_t = 0$)
c	the adiabatic speed of sound
ω	the angular frequency
k	the wave number ($k = \omega/c$)
k_a	the acoustic wave number
k_h	the entropic wave number
k_v	the viscous wave number
β	the increase in pressure per unit increase in temperature at constant density
μ	dynamic viscosity
η	bulk viscosity
ρ_0	the density
κ	thermal conductivity
C_p	the heat capacity per unit mass at constant pressure
C_v	the heat capacity per unit mass at constant volume
γ	specific heat ratio ($\gamma = C_p/C_v$)
$G(X,Y)$	free space Green's function of Helmholtz's equation
S_p	surface where the pressure is known
S_v	surface where the velocity is known
m	azimuthal rank
i	complex number ($i^2 = -1$)

INTRODUCTION

The study of acoustic propagation when taking into account the viscosity and thermal conductivity of a fluid has received a lot of interest during the last 20 years with the development of miniaturized transducers. When the size of the cavity is in the order of the thickness of the acoustic or thermal boundary layer, the inclusion of these effects becomes necessary.[1–6]

For simple geometry, an analytic solution can be established.[7–10] However, for complex geometries, the recourse to numerical solutions proves necessary. We can quote especially the work of Dokumaci[11] who investigates, using an integral equation formulation, the acoustic radiation problem in a viscous fluid without thermal conductivity.

This paper presents a variational formulation using integral equations for the resolution of the problem of propagation in a viscothermal fluid. This approach avoids, on one hand, the meshing of the acoustic medium and on the other hand, eliminates the problems of singularities linked to the representation by integral equations. After discretization by finite boundary elements, we obtain a linear symmetrical algebraic system of relatively small size. Numerical results predicted by this formulation have been compared to analytical results that are developed for a simple case (cylindrical cavity subjected to the mixed limit conditions).

I. EQUATIONS AND BOUNDARY CONDITIONS

The basic relations which govern the characteristic variables of a viscothermal fluid (Fig. 1) such as the particle velocity $\mathbf{v}$, the density fluctuation ρ, the pressure variation p, and the temperature fluctuation T, are obtained by a suitable decomposition of the Navier–Stokes equations, the continuity equation, the energy equation, and the thermodynamic state equations. In the absence of external forces and rate of heat creation and when at initial state the fluid is at rest (i.e., density, pressure, etc, are uniform in space and time), the set of linear homogeneous equations governing small amplitude disturbances for harmonic motion ($\partial_t = -i\omega$), can be easily demonstrated.[4–6]

The pressure variation p is the sum of an acoustic pressure p_a and entropic pressure p_h:

$$p = p_a + p_h, \tag{1}$$
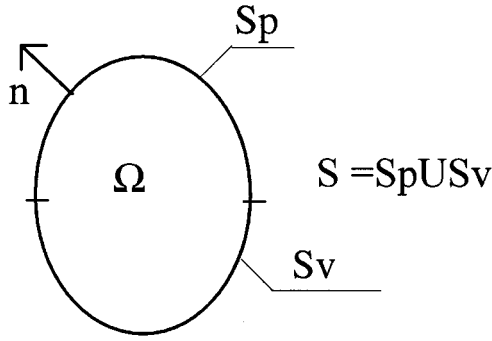
FIG. 1. Viscothermal fluid occupant the domain Ω.

$$(\Delta + k_a^2)p_a = 0, \tag{2}$$

$$(\Delta + k_h^2)p_h = 0, \tag{3}$$

where the two kinds of waves [Eqs. (2) and (3)] are known as propagational (or acoustic) and thermal (or entropy) mode.

The temperature fluctuation T is the sum of an acoustic temperature T_a and entropic temperature T_h related to the acoustic and entropic pressures by

$$T = T_a + T_h = \tau_a p_a + \tau_h p_h. \tag{4}$$

The velocity is the sum of a rotational velocity $\mathbf{v}_t$ and an irrotational velocity $\mathbf{v}_e$:

$$\mathbf{v} = \mathbf{v}_e + \mathbf{v}_t, \tag{5}$$

$$(\Delta + k_v^2)\mathbf{v}_t = 0, \quad \text{div } \mathbf{v}_t = 0, \tag{6}$$

$$\mathbf{v}_e = \mathbf{v}_e^a + \mathbf{v}_e^h = \phi_a \text{ grad } p_a + \phi_h \text{ grad } p_h, \tag{7}$$

where the kind of wave [Eq. (6)] is known as vorticity (or shear) mode.

In these equations we have

$$k_a^2 = k^2(1 - ikl_{vh} - k^2 l_h l'_{vh})^{-1}, \quad k_h^2 = \frac{ik}{l_h}(1 + ikl'_{vh})^{-1},$$

$$k_v^2 = \frac{ik}{l_v}, \quad \tau_a = \frac{\gamma - 1}{\beta\gamma}\left(1 + i\frac{c}{\omega}l_h k_a^2\right)^{-1},$$

$$\tau_h = \frac{\gamma - 1}{\beta\gamma}\left(1 + i\frac{c}{\omega}l_h k_h^2\right)^{-1},$$

$$\phi_a = \frac{-i}{\rho_0\omega}\left(1 + i\frac{c}{\omega}l'_v k_a^2\right)^{-1}, \quad \phi_h = \frac{-i}{\rho_0\omega}\left(1 + i\frac{c}{\omega}l'_v k_h^2\right)^{-1},$$

where l_h, l_v, l'_v, l_{vh}, and l'_{vh} are the characteristic lengths related to viscosity coefficients and thermal diffusivity as follows:

$$l_h = \frac{\kappa}{\rho_0 C_p c}, \quad l_v = \frac{\mu}{\rho_0 c}, \quad l'_v = \frac{\eta + 4/3\mu}{\rho_0 c},$$

$$l_{vh} = l'_v + (\gamma - 1)l_h, \quad l'_{vh} = (\gamma - 1)(l_h - l'_v).$$

To extend this formalism to the edge of domain Ω, we must add the boundary conditions which are of two kinds: thermal boundary conditions and acoustic boundary conditions. In most applications, it is well known that the product of the thermal conductivity and specific heat of walls enclosing the frontiers of domain Ω greatly exceeds the corresponding quantity of the fluid itself; so, the temperature fluctuation is null on the boundaries.[12–14] We consider that the edge of domain Ω is composed of two surfaces, S_p when a pressure $\bar{p}$ is imposed and S_v when a velocity $\bar{V}$ is imposed, so that we have

over S_p:

$$p_a = \frac{-\tau_h \bar{p}}{\tau_a - \tau_h}, \tag{8}$$

$$p_h = \frac{\tau_a \bar{p}}{\tau_a - \tau_h}. \tag{9}$$

over S_v:

$$\phi_a \frac{\partial p_a}{\partial n} + \phi_h \frac{\partial p_h}{\partial n} + \mathbf{v}_t^n = \bar{V}_n \quad \text{(normal component)}, \tag{10}$$

$$\phi_a \frac{\partial p_a}{\partial \tau} + \phi_h \frac{\partial p_h}{\partial \tau} + \mathbf{v}_t^\tau = \bar{V}_\tau \quad \text{(tangential component)}, \tag{11}$$

$$T = \tau_a p_a + \tau_h p_h = 0. \tag{12}$$

Also, it is assumed that the normal component of rotational velocity $\mathbf{v}_t$ is negligible as compared to the irrotational velocity $\mathbf{v}_e$,[11] therefore, Eq. (10) is simplified as

$$\phi_a \frac{\partial p_a}{\partial n} + \phi_h \frac{\partial p_h}{\partial n} = \bar{V}_n. \tag{13}$$

Thus we decoupled the problem of pressure calculation as well as that of rotational velocity. Note that we are going to solve only the former.

II. INTEGRAL REPRESENTATION OF ACOUSTIC AND ENTROPIC PRESSURES

An integral representation of acoustic and entropic pressures is used for solving the pressure calculation problem. It is based on determining the two potentials,[15] one of simple layer and the other of double layer:

$$p_a(X) = \int_S \left(-p_{aY}\frac{\partial G_a}{\partial n_Y} + G_a(X,Y)\frac{\partial p_a}{\partial n_Y}\right)dS_Y,$$

$$X \in \Omega/S \tag{14}$$

$$p_h(X) = \int_S \left(-p_{hY}\frac{\partial G_h}{\partial n_Y} + G_h(X,Y)\frac{\partial p_h}{\partial n_Y}\right)dS_Y,$$

$$X \in \Omega/S, \tag{15}$$

where p_{aY} is the acoustic pressure at point Y which belongs to the surface S, $\mathbf{n}$ is the external normal vector of regions containing fluid; G_a and G_h are the elementary solutions of Eqs. (2) and (3) in the free space.

Considering the boundary conditions in the above equations gives

$$p_a(X) = \int_{S_p} \left(\frac{\tau_h}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial G_a}{\partial n_Y} + G_a \frac{\partial p_a}{\partial n_Y} \right) dS_Y$$

$$+ \int_{S_v} \left(G_a \frac{\partial p_a}{\partial n_Y} - p_{aY} \frac{\partial G_a}{\partial n_Y} \right) dS_Y, \tag{16}$$

$$p_h(X) = \int_{S_p} \left(\frac{-\tau_a}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial G_h}{\partial n_Y} + G_h \frac{\partial p_h}{\partial n_Y} \right) dS_Y$$

$$+ \int_{S_v} \left[\frac{\tau_a}{\tau_h} p_{aY} \frac{\partial G_h}{\partial n_Y} + \phi_h^{-1} G_h \left(\bar{V}_{nY} \right. \right.$$

$$\left. \left. - \phi_a \frac{\partial p_a}{\partial n_Y} \right) \right] dS_Y, \quad X \in \Omega/S. \tag{17}$$

The application of boundary conditions (8) and (9) over S_p, (12) and (13) over S_v, at integral representations of p_a and p_h and their normal derivatives, gives the system of equations which enables us to determine the unknowns of the problem:
$\partial p_a/\partial n$ and $\partial p_h/\partial n$ over S_p, p_a and $\partial p_a/\partial n$ over S_v.
For $X \in S_p$:

$$\int_{S_p} G_a \frac{\partial p_a}{\partial n_Y} dS_Y - \int_{S_v} p_{aY} \frac{\partial G_a}{\partial n_Y} dS_Y + \int_{S_v} G_a \frac{\partial p_a}{\partial n_Y} dS_Y$$

$$= F_1(X), \tag{18}$$

where

$$F_1(X) = -\text{p.v.} \int_{S_p} \left(\frac{\tau_h}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial G_a}{\partial n_Y} \right) dS_Y - \frac{\tau_h}{\tau_a - \tau_h} \frac{\bar{p}_X}{2},$$

$$\int_{S_p} G_h \frac{\partial p_h}{\partial n_Y} dS_Y + \int_{S_v} \frac{\tau_a}{\tau_h} p_{aY} \frac{\partial G_h}{\partial n_Y} dS_Y$$

$$- \int_{S_v} \phi_h^{-1} \phi_a G_h \frac{\partial p_a}{\partial n_Y} = F_2(X), \tag{19}$$

where

$$F_2(X) = \frac{\tau_a}{\tau_a - \tau_h} \frac{\bar{p}_X}{2} + \text{p.v.} \int_{S_p} \frac{\tau_a}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial G_h}{\partial n_Y}$$

$$- \phi_h^{-1} \int_{S_v} G_h \bar{V}_{nY} dS_Y.$$

For $X \in S_v$:

$$\phi_a \int_{S_p} \left(\frac{\partial G_a}{\partial n_X} \frac{\partial p_a}{\partial n_Y} \right) dS_Y + \phi_h \int_{S_p} \left(\frac{\partial G_h}{\partial n_X} \frac{\partial p_h}{\partial n_Y} \right) dS_Y$$

$$+ \phi_h \text{F.P.} \int_{S_v} \frac{\tau_a}{\tau_h} p_{aY} \frac{\partial^2 G_h}{\partial n_X \partial n_Y} dS_Y$$

$$- \phi_a \text{F.P.} \int_{S_v} p_{aY} \frac{\partial^2 G_a}{\partial n_X \partial n_Y} dS_Y + \phi_a \text{ p.v.}$$

$$\times \int_{S_v} \left(\frac{\partial G_a}{\partial n_X} - \frac{\partial G_h}{\partial n_X} \right) \frac{\partial p_a}{\partial n_Y} dS_Y = F_3(X), \tag{20}$$

where

$$F_3(X) = \frac{\bar{V}_{nX}}{2} - \phi_a \int_{S_p} \left(\frac{\tau_h}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial^2 G_a}{\partial n_X \partial n_Y} \right) dS_Y$$

$$+ \phi_h \int_{S_p} \left(\frac{\tau_a}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial^2 G_h}{\partial n_X \partial n_Y} \right) dS_Y$$

$$- \text{p.v.} \int_{S_v} \left(\bar{V}_{nY} \frac{\partial G_h}{\partial n_X} \right) dS_Y,$$

$$\tau_a \int_{S_p} G_a \frac{\partial p_a}{\partial n_Y} dS_Y + \tau_h \int_{S_p} G_h \frac{\partial p_h}{\partial n_Y} dS_Y$$

$$+ \tau_a \text{ p.v.} \int_{S_v} \left(\frac{\partial G_h}{\partial n_Y} - \frac{\partial G_a}{\partial n_Y} \right) p_{aY} dS_Y$$

$$+ \int_{S_v} (\tau_a G_a - \tau_h \phi_h^{-1} \phi_a G_h) \frac{\partial p_a}{\partial n_Y} dS_Y = F_4(X), \tag{21}$$

where

$$F_4(X) = \int_{S_p} \left(\frac{-\tau_a \tau_h}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial G_a}{\partial n_Y} \right) dS_Y + \int_{S_p} \left(\frac{\tau_h \tau_a}{\tau_a - \tau_h} \bar{p}_Y \frac{\partial G_h}{\partial n_Y} \right)$$

$$\times dS_Y - \tau_h \int_{S_v} \phi_h^{-1} G_h \bar{V}_{nY} dS_Y$$

and where p.v. is the principal value of Cauchy and F.P. is the finite part of Hadamard.

Therefore a system of four equations [(18), (19), (20), and (21)] with four unknowns: $\partial p_a/\partial n$, $\partial p_h/\partial n$ over S_p; p_a, $\partial p_a/\partial n$ over S_v is obtained. This system can be solved directly using, for example, the method of collocation.[16] However, this method has two disadvantages: First, numerical computation of the ''finite part'' of the singular integrals and second, the final algebraic system is nonsymmetrical. To overcome these disadvantages, we suggest to associate a variational formulation to this system. With this variational formulation, the explicit evaluation of the ''finite part'' of the singular integrals is avoided and the final system is symmetric.

III. VARIATIONAL FORMULATION

Let $(\partial p_a/\partial n)'(S_p)$, $(\partial p_h/\partial n)'(S_p)$, $p_a'(S_v)$, and $(\partial p_a/\partial n)'(S_v)$ represent four sufficiently regular test functions $[(\partial p_a/\partial n)'(S_p)$ means the normal derivative of acoustic pressure at the surface $S_p]$. After multiplying Eqs. (18) and (19) with variables $(\partial p_a/\partial n)'$ and $(\partial p_h/\partial n)'$ and inte-

1313   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

C. Karra and M. Ben Tahar: Viscothermal fluids   1313

grating over S_p, also Eqs. (20) and (21) with variables p'_a and $(\partial p_a/\partial n)'$ and integrating over S_v, making an appropriate linear combination for the purpose to have a symmetrical variational formulation gives

$$\langle \cdot \rangle_{S_p \times S_p} + \langle \cdot \rangle_{S_v \times S_p} + \langle \cdot \rangle_{S_p \times S_v} + \langle \cdot \rangle_{S_v \times S_v}$$

$$= \left\langle F_1(X), \left(\frac{\partial p_a}{\partial n_X}\right)' \right\rangle_{S_p} - \frac{\tau_h}{\tau_a} \frac{\phi_h}{\phi_a} \left\langle F_2(X), \left(\frac{\partial p_h}{\partial n_X}\right)' \right\rangle_{S_p}$$

$$- \frac{1}{\phi_a} \langle F_3(X), p'_{aX} \rangle_{S_v} + \frac{1}{\tau_a} \left\langle F_4(X), \left(\frac{\partial p_a}{\partial n_X}\right)' \right\rangle_{S_v},$$

$$(22)$$

where the double integrals $\langle \cdot \rangle$ are presented in Appendix A.

IV. DISCRETIZATION BY FINITE BOUNDARY ELEMENTS METHOD

For the discretization of variational formulation (22), we consider problems with revolution geometry. In cylindrical coordinate system (r,z,θ), the Fourier series decomposition of fluid variables Ψ:

$$\Psi(r,z,\theta) = \sum_{m=-\infty}^{m=+\infty} \Psi_m(r,z) e^{im\theta} \quad (23)$$

enables us to convert a three-dimensional problem (r,z,θ) into a series of two-dimensional problem (r,z). This transition requires the calculation of Fourier components of Green's function G and its first derivative. These components have been developed by Mebarek.[17] The solution of the three-dimensional problem is then obtained by summing according to (23) the solution of each modal problem.

For the discretization, we use a linear isoparametric element with two nodes and two degrees of freedom. These are $(\partial p_a/\partial n, \partial p_h/\partial n)$ if the node belongs to a surface S_p or $(\partial p_a/\partial n, p_a)$ if the node belongs to a surface S_v.

Linear shape functions are used to express the coordinates of a point X as a function of nodes coordinates $(X1,X2)$ and fluid variables Ψ as a function of nodal variables $(\Psi 1, \Psi 2)$. The discretization by finite boundary elements of the variational formulation (22) for a given azimuthal mode m, permits to obtain after assembly the symmetrical matrix system

$$\begin{bmatrix} A_1 & 0 & \vdots & A_6 & A_8 \\ & A_2 & \vdots & A_7 & A_9 \\ \hdashline & & \vdots & A_3 & A_5 \\ & & \vdots & & A_4 \end{bmatrix} \begin{Bmatrix} \frac{\partial p_a}{\partial n}(S_p) \\ \frac{\partial p_h}{\partial n}(S_p) \\ \frac{\partial p_a}{\partial n}(S_v) \\ p_a(S_v) \end{Bmatrix} = \begin{Bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{Bmatrix}, \quad (24)$$

where $\{Si\}$ denotes the discretised second vector due to acoustic excitations $\bar{p}$ and $\bar{V}$.
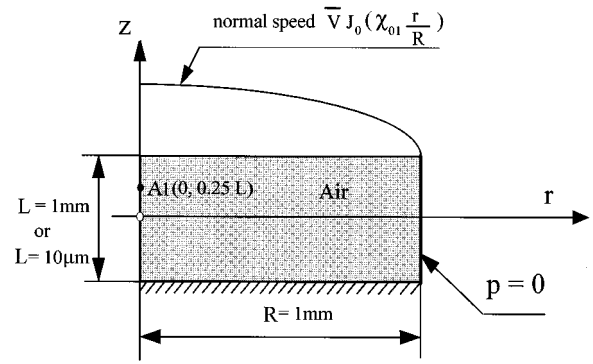
FIG. 2. Cylindrical cavity.

V. ANALYTICAL AND NUMERICAL RESULTS

In this section we present an example for validating numerical developments. It concerns a cylindrical cavity filled with air, supposed to be viscothermal (Fig. 2) whose physical properties are: $\rho_0 = 1.1614$ kg/m^3, $c = 340$ m/s, $\mu = 184.6 \times 10^{-7}$ Ns/m^2, $\eta = 110 \times 10^{-7}$ Ns/m^2, $\beta = 458$ N/(m^2 K), $\chi = 22.5 \times 10^{-6}$ m^2/s, $\gamma = 1.403$. Numerical viscothermal code results are compared with the analytic solution.

The temperature fluctuation is assumed to be null at all boundaries. The pressure is zero at the boundary $r=R$. For $z=L/2$, the normal speed is imposed to $\bar{V}J_0[\chi_{01}(r/R)]$ where $\bar{V}=1$ m/s (one considers the azimuthal mode $m=0$) and χ_{01} is the first zero of the Bessel function J_0. Finally, at $z=-L/2$, the speed is null.

Acoustic and entropic pressure solutions of Eqs. (2) and (3) which verify the above boundary conditions are

$$p_a = J_0\left(\frac{\chi_{01}r}{R}\right)\left[A_{a1} \exp\left(-ik_{az1}\left(z-\frac{L}{2}\right)\right) \right.$$

$$\left. + B_{a1} \exp\left(ik_{az1}\left(z+\frac{L}{2}\right)\right)\right], \quad (25)$$

$$p_h = J_0\left(\frac{\chi_{01}r}{R}\right)\left[A_{h1} \exp\left(-ik_{hz1}\left(z-\frac{L}{2}\right)\right) \right.$$

$$\left. + B_{h1} \exp\left(ik_{hz1}\left(z+\frac{L}{2}\right)\right)\right], \quad (26)$$

where $k_{az1}^2 = k_a^2 - \chi_{01}^2/R^2$, $k_{hz1}^2 = k_h^2 - \chi_{01}^2/R^2$ and A_{a1}, A_{h1}, B_{a1}, and B_{h1} are determined from boundary conditions (see Appendix B).

In the particular case where the air is considered perfect and a nonthermal conductor, the pressure in the cavity is given by

$$p = J_0\left(\frac{\chi_{01}r}{R}\right)\left[A \exp\left(-ik_z\left(z-\frac{L}{2}\right)\right) \right.$$

$$\left. + B \exp\left(ik_z\left(z+\frac{L}{2}\right)\right)\right], \quad (27)$$

where
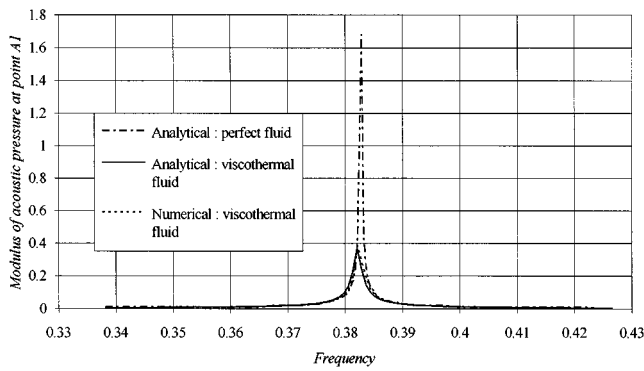
FIG. 3. Modulus of the acoustic pressure as a function of frequency.

$$k_z = \sqrt{k^2 - \frac{\chi_{01}^2}{R^2}} \quad \text{if} \quad k \geqslant \frac{\chi_{01}}{R},$$

$$k_z = +i\sqrt{\frac{\chi_{01}^2}{R^2} - k^2} \quad \text{if} \quad k < \frac{\chi_{01}}{R},$$

$$A = \frac{-\rho_0 \omega \overline{V}}{2k_z}\left(\frac{1}{1 - \exp(ik_z L)} + \frac{1}{1 + \exp(ik_z L)}\right),$$

$$B = \frac{-\rho_0 \omega \overline{V}}{2k_z}\left(\frac{1}{1 - \exp(ik_z L)} - \frac{1}{1 + \exp(ik_z L)}\right).$$

Natural frequencies of the cavity are given by

$$f_{mnl} = \frac{c}{2\pi}\sqrt{\frac{\chi_{mn}^2}{R^2} + \frac{\pi^2 l^2}{L^2}}, \quad m,l = 0,1,\dots, \quad n = 1,2,\dots \,.$$

All analytical and numerical results presented in this paper are nondimensionalized by referring them to the following parameters: ρ_0 (the density), c (the adiabatic speed of sound), and L (the height of the cavity).
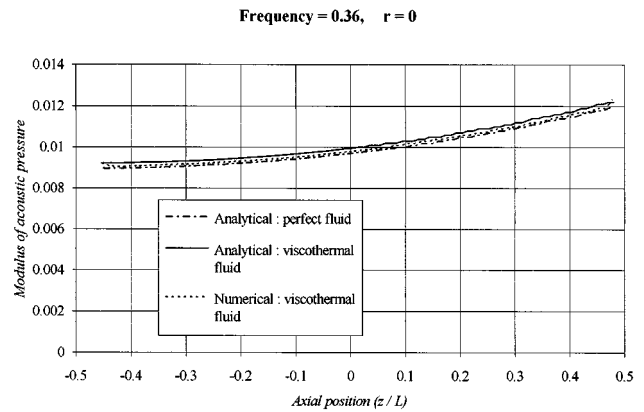
FIG. 5. Modulus of acoustic pressure as a function of axial position for the nondimensional frequency 0.36.
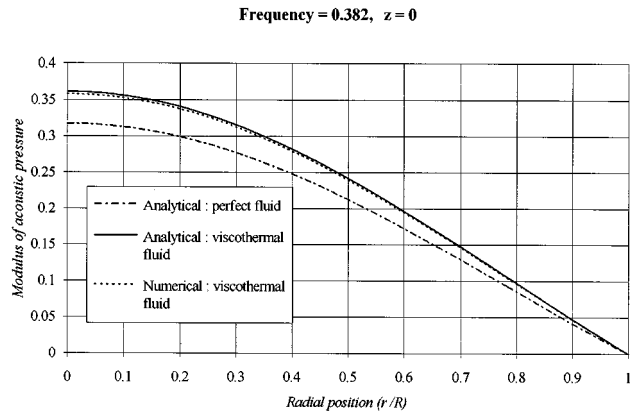
FIG. 6. Modulus of the acoustic pressure as a function of radial position for the nondimensional frequency 0.382.
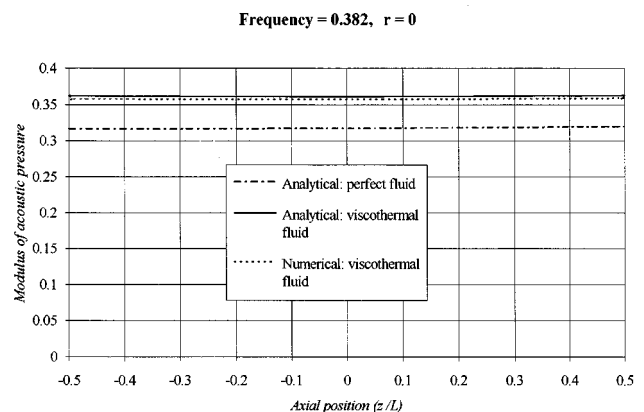
FIG. 4. Modulus of the acoustic pressure as a function of radial position for the nondimensional frequency 0.36.

FIG. 7. Modulus of acoustic pressure as a function of axial position for the nondimensional frequency 0.382.

1315   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

C. Karra and M. Ben Tahar: Viscothermal fluids   1315

FIG. 8. Modulus of the entropic pressure as a function of axial position for the nondimensional frequency 0.36.

FIG. 9. Modulus of the acoustic pressure as a function of frequency.
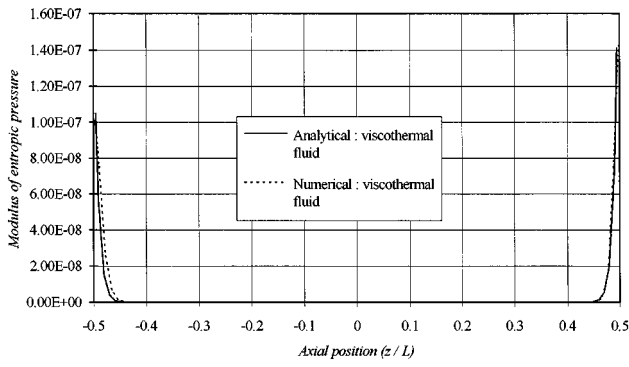
FIG. 10. Modulus of the acoustic pressure as a function of radial position for the nondimensional frequency 0.0033.
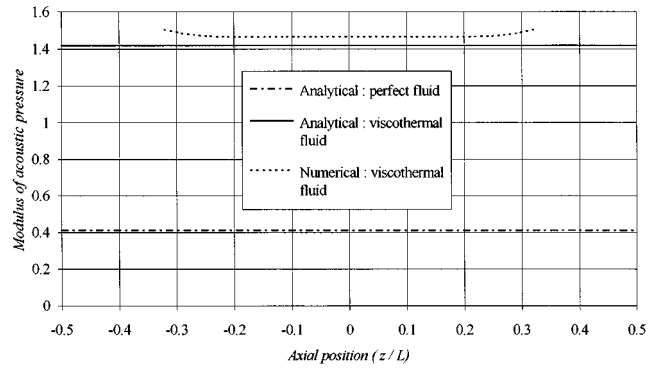
FIG. 11. Modulus of acoustic pressure as a function of axial position for the nondimensional frequency 0.0033.
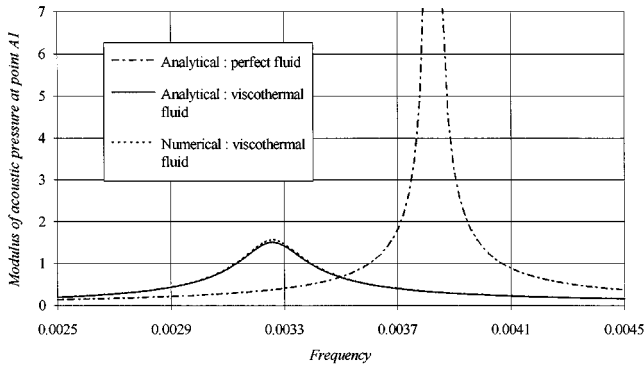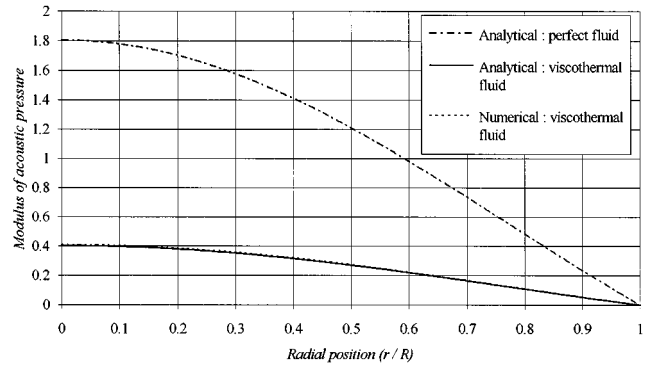
FIG. 12. Modulus of the acoustic pressure as a function of radial position for the nondimensional frequency 0.0037.

FIG. 13. Modulus of acoustic pressure as a function of axial position for the nondimensional frequency 0.0037.

Frequency = 0.0037, r = 0
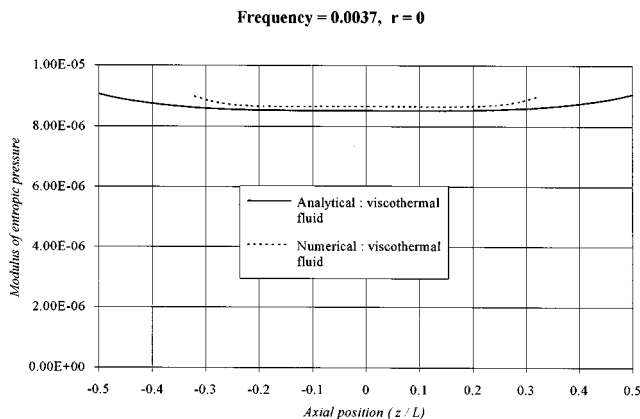
FIG. 14. Modulus of the entropic pressure as a function of axial position for the nondimensional frequency 0.0037.

A. Cavity with the height of 1 mm

Figure 3 shows the nondimensional modulus of the acoustic pressure at point Al (Fig. 2), versus the nondimensional frequency. The obtained peak corresponds to the first radial mode ($f = 0.3827$). Numerical and analytical results of the viscothermal model are in perfect agreement. In addition, one sees well the effects of air viscosity on the damping of the first radial mode.

For the nondimensional frequency 0.36, we have carried the modulus of the acoustic pressure according to the radial position ($z = 0$) in Fig. 4 and to the axial position ($r = 0$) in Fig. 5. We observe that numerical and analytical results of the viscothermal model are in perfect agreement. They are close to those of the perfect fluid model, since this frequency is far from the resonance of the first mode. However, for the nondimensional frequency 0.382 (very close to the first resonance frequency of the cavity), Figs. 6 and 7 show that the acoustic pressure of the viscothermal fluid model is different from that of the perfect fluid model.

The nondimensional modulus of the entropic pressure versus the axial position is reported (Fig. 8) at the nondimensional frequency 0.36. Numerical results are very close to the analytical ones. This shows the existence of two thermal boundary layers at the boundaries ($z = \pm L/2$). Indeed, thermal waves having a wave number $k_h = (i+1)\sqrt{k/2l_h}$ undergoes an intense damping, thereby from their origin these waves exist only in a region of small thickness called the thermal boundary layer. The thickness of this boundary layer corresponds to the distance d_h at the end of which the wave is damped by the ratio $e = 2.718$. For the nondimensional frequency 0.36, the real thickness of the thermal boundary layer is $d_h = \sqrt{2l_h c/\omega} = 7.65$ μm.

By referring to the results of Fig. 8, we deduce that the distance at the end of which the amplitude of the entropic pressure is damped by the ratio e is $d_h \cong 7.5$ μm, it is therefore close to the value calculated.

B. Cavity with the height of 10 μm

The same results have been obtained for a case of a cavity with the height of 10 μm and numerical results compared to analytic ones. Figure 9 shows that the numerical and analytic results of the viscothermal model are in good agreement and are clearly different from those of the perfect fluid model. One clearly sees the damping of the first radial mode and the decrease of the first resonance frequency which is about ($\sim 15\%$).

For two particular frequencies 0.0033 (close to the first resonance frequency of the cavity with viscothermal model) and 0.0037 (close to the first resonance frequency of the cavity with perfect fluid model), we have carried the nondimensional modulus of acoustic pressure according to radial position ($z = 0$) and axial position ($r = 0$) in Figs. 10–13. These figures show the validity of our numerical results and illustrate clearly the difference between the results of the perfect fluid model and those of the viscothermal model.

We also computed the nondimensional modulus of entropic pressure for the frequency 0.0037 according to axial position in Fig. 14. Here, we cannot illustrate clearly like in Fig. 8 (which concerns cavity with the height of 1 mm) the two thermal boundary layers at the boundaries ($z = \pm L/2$), because in this case these thermal boundary layers filled up the whole cavity. Therefore, the entropic pressure is practically uniform through the height of the microcavity.

VI. CONCLUSION

An original variational formulation by integral equations for the resolution of an acoustic propagation problem in a viscothermal fluid is presented. This formulation allows us to take into account complex geometry and multiple acoustic boundary conditions.

The numerical implementation using finite boundary elements of this formulation for problems of geometry of revolution has been validated analytically.

We clearly show the effects of air viscosity on the damping of the first radial mode of the cavity as well as thermal effects at the boundaries. The results illustrate these effects especially in the case of microcavities. The good representation of these effects on the edges of the cylinder allows to better apprehend thereafter the vibroacoustic problem.

ACKNOWLEDGMENTS

APPENDIX A

The integrals $\langle \cdot \rangle$ presented in the first member of variational formulation (27) are given by

$$\langle \cdot \rangle_{S_p \times S_p} = \left\langle G_a \frac{\partial p_a}{\partial n_Y} \left(\frac{\partial p_a}{\partial n_X} \right)' \right.$$

$$\left. - \frac{\tau_h}{\tau_a} \frac{\phi_h}{\phi_a} G_h \frac{\partial p_h}{\partial n_Y} \left(\frac{\partial p_h}{\partial n_X} \right)' \right\rangle_{S_p \times S_p},$$

$$\langle\,\cdot\,\rangle_{\mathbf{S_v}\times\mathbf{S_v}}=\left\langle\left(\frac{\partial^2 G_a}{\partial n_X\partial n_Y}-\frac{\phi_h}{\phi_a}\frac{\tau_a}{\tau_h}\frac{\partial^2 G_h}{\partial n_X\,\partial n_Y}\right)p_{aY}p'_{aX}\right\rangle_{\mathbf{S_v}\times\mathbf{S_v}}$$

$$+\left\langle\left(\left(\frac{\partial G_h}{\partial n_X}-\frac{\partial G_a}{\partial n_X}\right)p'_{aX}\frac{\partial p_a}{\partial n_Y}\right.\right.$$

$$\left.+\left(\frac{\partial G_h}{\partial n_Y}-\frac{\partial G_a}{\partial n_Y}\right)p_{aY}\left(\frac{\partial p_a}{\partial n_X}\right)'\right\rangle_{\mathbf{S_v}\times\mathbf{S_v}}$$

$$+\left\langle\left(\left(G_a-\frac{\tau_h}{\tau_a}\frac{\phi_a}{\phi_h}G_h\right)\frac{\partial p_a}{\partial n_Y}\left(\frac{\partial p_a}{\partial n_X}\right)'\right\rangle_{\mathbf{S_v}\times\mathbf{S_v}},$$

$$\langle\,\cdot\,\rangle_{\mathbf{S_v}\times\mathbf{S_p}}=\left\langle-\frac{\partial G_a}{\partial n_Y}\,p_{aY}\left(\frac{\partial p_a}{\partial n_X}\right)'\right\rangle_{\mathbf{S_v}\times\mathbf{S_p}}$$

$$+\left\langle-\frac{\phi_h}{\phi_a}\frac{\partial G_h}{\partial n_Y}\,p_{aY}\left(\frac{\partial p_h}{\partial n_X}\right)'\right\rangle_{\mathbf{S_v}\times\mathbf{S_p}}$$

$$+\left\langle\frac{\tau_h}{\tau_a}\,G_h\,\frac{\partial p_a}{\partial n_Y}\left(\frac{\partial p_h}{\partial n_X}\right)'\right\rangle_{\mathbf{S_v}\times\mathbf{S_p}}$$

$$+\left\langle G_a\,\frac{\partial p_a}{\partial n_Y}\left(\frac{\partial p_a}{\partial n_X}\right)'\right\rangle_{\mathbf{S_v}\times\mathbf{S_p}},$$

$$\langle\,\cdot\,\rangle_{\mathbf{S_p}\times\mathbf{S_v}}=\left\langle-\frac{\partial G_a}{\partial n_X}\frac{\partial p_a}{\partial n_Y}\,p'_{aX}\right\rangle_{\mathbf{S_p}\times\mathbf{S_v}}$$

$$+\left\langle-\frac{\phi_h}{\phi_a}\frac{\partial G_h}{\partial n_X}\frac{\partial p_h}{\partial n_Y}\,p'_{aX}\right\rangle_{\mathbf{S_p}\times\mathbf{S_v}}$$

$$+\left\langle\frac{\tau_h}{\tau_a}\,G_h\,\frac{\partial p_h}{\partial n_Y}\left(\frac{\partial p_a}{\partial n_X}\right)'\right\rangle_{\mathbf{S_p}\times\mathbf{S_v}}$$

$$+\left\langle G_a\,\frac{\partial p_a}{\partial n_Y}\left(\frac{\partial p_a}{\partial n_X}\right)'\right\rangle_{\mathbf{S_p}\times\mathbf{S_v}},$$

$$\langle\,\cdot\,\rangle_{\mathbf{S_p}\times\mathbf{S_v}}=\int_{\mathbf{S_v}}\left(\int_{\mathbf{S_p}}\cdot\,dS_Y\right)dS_X.$$

APPENDIX B

The four constants A_{a1}, A_{h1}, B_{a1}, and B_{h1} presented in the Eqs. (25) and (26) are determined by solving the following four equations:

$$-i\{k_{az1}\phi_a[A_{a1}-B_{a1}\,\exp(ik_{az1}L)]$$

$$+k_{hz1}\phi_h[A_{h1}-B_{h1}\,\exp(ik_{hz1}L)]\}=\bar{V},$$

$$\tau_a[A_{a1}+B_{a1}\,\exp(ik_{az1}L)]+\tau_h[A_{h1}+B_{h1}\,\exp(ik_{hz1}L)]=0,$$

$$-i\{k_{az1}\phi_a[A_{a1}\,\exp(ik_{az1}L)-B_{a1}]$$

$$+k_{hz1}\phi_h[A_{h1}\,\exp(ik_{hz1}L)-B_{h1}]\}=0,$$

$$\tau_a[A_{a1}\,\exp(ik_{az1}L)+B_{a1}]+\tau_h[A_{h1}\,\exp(ik_{hz1}L)+B_{h1}]=0.$$

[1] J. P. Lefebvre, ''Petites Perturbations d'un Fluide Visqueux Conducteur de la Chaleur dans un état Initial Quasi-stationnaire et Quasi-uniforme. Opérateur de Diffusion-propagation et Application à l'acoustique Linéaire,'' Rev. Cethedec **51**, 103–120 (1977).

[2] B. Poirée, ''Les ondes planes hétérogènes dans un fluide thermovisqueux,'' Acustica **64**, 73–79 (1987).

[3] M. Dechamps, ''L'équation de dispersion des ondes thermiques en milieu fluide thermovisqueux,'' C. R. Acad. Sci. Paris, Ser. II **308**, 599–602 (1989).

[4] P. M. Morse and K. Uno Ingard, *Theoretical Acoustics* (Princeton U.P., Princeton, NJ, 1986).

[5] M. Bruneau, *Introduction aux théories de l'acoustique* (Université du Maine, Le Mans, France, 1983).

[6] M. Bruneau, Ph. Herzog, J. Kergomard, and J. D. Polack, ''General formulation of dispersion equation in bounded visco-thermal fluid, and application to simple geometries,'' Wave Motion **11**, 441–451 (1989).

[7] M. Bruneau and G. Plantier, ''Heat conduction effects on the acoustic response of a membrane separated by a very thin air film from a backing electrode,'' J. Acoust. **3**, 243–250 (1990).

[8] M. Bruneau, ''Acoustics fluid-filled small cavities,'' Proc. Inst. Acoust. **15**, 1–20 (1993).

[9] M. Bruneau, A. M. Bruneau, and P. Hamery, ''Modélisation des microphones miniatures: Effets des couches limites viscothermiques,'' J. Phys. (France) IV, Colloq. C1, Suplément au J. Phys. III **2**, 361–364 (1992).

[10] M. Bruneau and A. M. Bruneau, ''Dynamic behaviour of a thin fluid layer coupled to a membrane,'' C. R. Acad. Sci. Paris **317**, Ser II, 1543–1548 (1993).

[11] E. Dokumaci, ''An integral equation formulation for boundary element analysis of acoustic radiation problems in viscous fluids,'' J. Sound Vib. **147**, 335–348 (1991).

[12] K. Rathnam, ''Influence of velocity slip and temperature jump in rarefied gaz acoustic oscillations in cylindrical tubes,'' J. Sound Vib. **103**, 448–452 (1985).

[13] A. Tominaga, ''Thermoacoustic effects of invicid fluids,'' J. Low Temp. Phys. **54**, 233–245 (1984).

[14] LM. J. H. Fox, ''The damping of structural vibration by thin gas films,'' J. Sound Vib. **73**, 279–295 (1980).

[15] M. A. Hamdi, ''Formulation variationnelle par équations intégrales pour la résolution de l'équation de Helmholtz avec des conditions aux limites mixtes,'' C. R. Acad. Sci. Paris, Ser. II **292**, 17–20 (1981).

[16] M. N. Sayhi, Y. Ousset, and G. Verchery, ''Solution of radiation problem by collocation of integral formulation in terms of single and double layer potentials,'' J. Sound Vib. **74**, 187–204 (1981).

[17] L. Mebarek, ''Méthode numérique par éléments finis de frontières pour le calcul par équations intégrales de champs acoustiques rayonnés par des structures axisymétriques,'' Thèse de doctorat de l'Université de Technologie de Compiègne, France, 1986.

Complex multipole beam approach to acoustic scattering problems

Amir Boag
Israel Aircraft Industries, Department 4464, Ben-Gurion Airport 70100, Israel

Alona Boag
16 Erez Street, Yavne 81510, Israel

Eric Michielssen
Department of Electrical and Computer Engineering, University of Illinois, Urbana, Illinois 61801

A novel approach to reducing the matrix size associated with the boundary element/moment method solution of the problem of acoustic scattering from arbitrarily shaped closed bodies is presented in this paper. The key step in this approach is to represent the scattered field in terms of a series of beams produced by multipole sources located in a *complex* space. On the scatterer boundary, the fields generated by these multipole sources resemble the Gabor basis functions. By utilizing the properties of the Gabor series, guidelines for selecting the orders as well as locations of the multipole sources are developed. The present approach not only reduces the number of unknowns, but also generates a generalized impedance matrix with a banded structure. The accuracy of the proposed method is verified by using internal accuracy checks and by comparing the numerical results with the analytic solutions for spherical scatterers. © *1997 Acoustical Society of America.* [S0001-4966(97)02309-6]

PACS numbers: 43.20.Fn, 43.20.Gp, 43.20.Bi [JEG]

INTRODUCTION

The conventional approach to formulating the problem of acoustic scattering using the boundary element/moment method (BE/MM) is to employ subdomain basis functions to represent the surface fields or equivalent sources on the body.[1,2] This approach often requires the use of 10 or more unknowns per linear wavelength or, equivalently, more than 100 unknowns per square wavelength, and thus leads to relatively large matrix sizes even for moderately sized scatterers. In addition, the resulting generalized impedance matrix is typically dense, owing to a strong coupling between the subdomain basis and testing functions, even when they are widely separated from each other. However, it is well known from asymptotic solutions of canonical problems, that the wave interactions on the smooth convex parts of the scatterer are essentially local. Canning[3–5] has shown that, for smooth scatterers, the use of Gabor type of basis functions,[6] *viz.*, functions comprised of windowed exponentials, can introduce the desired interaction/impedance matrix localization (IML) feature in the matrix equation. It should be noted, however, that the localization of the matrix elements near the diagonal is achieved in this approach at the expense of higher computational cost, introduced by the need to compute the double integrals for the matrix elements that involve the special basis and testing functions. The coupling between the distant parts of the scatterer due to the bi-directional character of the radiation emanating from the Gabor type of distributed sources is eliminated in the combined source IML approach.[4] One should note that Refs. 3 and 4 considered only two-dimensional problems, while in Ref. 5 IML has been extended to scattering by bodies of revolution.

Two alternate approaches, also proposed recently, are the multiple multipole (MMP)[7,8] and the source model[9,10] methods, both of which exhibit the potential for substantial savings in terms of the number of unknowns relative to the BE/MM. They employ the fields generated by fictitious sources, displaced from the boundary of the scatterer, to represent the scattered field. Both MMP and source model employ isotropic elemental sources, because the fields radiated by these sources can be conveniently expressed in analytical form. Use of such sources located away from the physical boundary results in a matrix that is not dominated by the diagonal. Recently, Pogorzelski[11] suggested the use of synthetic directive sources to represent the scattered field. The possibility of combining isotropic and directive sources has been examined by Erez and Leviatan.[12] While in Ref. 11 the 2D directive source is synthesized by an array of isotropic ones, in Ref. 12 the directivity is achieved by shifting a point source into complex space.[13] However, in both Ref. 11 and Ref. 12, only one type of directive source is used. This appears to be similar to Canning's first attempts to achieve localization employing only one type of basis function as described in Ref. 3. Thus, with the methods of Ref. 11 and Ref. 12 one can expect to achieve localization but, probably, at the expense of solving a highly ill-conditioned system of equations.

The objective of this paper is to apply to acoustic scattering the complex multipole beam approach (CMBA), which attempts to combine the advantages of both the IML and the MMP methods. The strategy followed in this method is to expand the scattered fields in terms of beams, generated by a judiciously selected set of multipole sources located in the complex space. These beams are very similar to the Gabor basis functions when sampled at the boundary of the

scatterer; hence, the method can be viewed as a numerical approach to finding an approximate Gabor representation of the boundary field. In fact, the problem of radiation from planar apertures has been extensively studied using the Gabor expansion of the aperture fields in two- and three-dimensional (2D and 3D) problems.[14,15] However, a straightforward Gabor expansion of the field over a nonplanar boundary would be of little use since the subsequent determination of the scattered field would be as difficult as solving the original problem itself. The completeness properties and other characteristics of the Gabor expansion functions are well understood,[16,17] and this greatly facilitates the task of developing a set of simple rules for choosing the orders and locations of the multipoles. And yet, in common with the MMP and the source model methods, the present approach retains the advantage in terms of the number of unknowns over the BE/MM, as it typically uses less than 4 unknowns per linear wavelength, or, equivalently, less than 15 unknowns per square wavelength.

The unidirectional character of the fields radiated by the multipole sources, located in the complex space, is similar to the nature of the combined electric and magnetic current basis functions of the Gabor type, used in the combined source IML approach.[4] The unidirectional nature of the source fields insures that the coupling will be low between a given source and all distant parts of the scatterer, including the part of the surface located on the opposite side. It also helps to suppress the spurious interior resonance type of solutions, since the fields generated inside the scatterer are relatively small.

It is worth mentioning that the analytic continuation of the multipole fields[18] to complex source coordinates is facilitated in the proposed method by expressing these fields in terms of Cartesian coordinates. This procedure is different from the one proposed by Shin and Felsen,[19] where they obtained the higher-order multipole beams by successively differentiating the zero-order beam along two transverse directions normal to the beam, a procedure which leads to only a subset of all multipoles with a standing-wave type of transverse behavior.

The three-dimensional (3D) acoustic formulation of the proposed method presented in this paper constitutes an extension of the 3D electromagnetic formulation presented in Ref. 20. Note that two-dimensional (2D) acoustic and electromagnetic formulations are equivalent. Thus, 2D CMBA, applied in Ref. 21 to the transverse magnetic (TM) scattering case, can be used (with mere substitution of appropriate quantities) to treat acoustic scattering from perfectly soft cylinders. In its present form, the method is most suitable for the analysis of large smooth bodies. A hybrid approach to extending the present method to a more general class of scatterers will be dealt with in forthcoming publications.

The organization of the paper is as follows. The properties of beams generated by multipoles located in the complex space and their relation to the Gabor expansion are discussed in the next section, i.e., Sec. I. The formulation of the method is outlined in Sec. II. Details of the numerical implementation of the method and illustrative results are presented in Sec. III. Finally, a few concluding remarks that summarize the paper are given in Sec. IV.

I. 3D MULTIPOLE BEAMS AND 2D GABOR EXPANSION

This section will be devoted to the study of the acoustic fields produced by a certain class of multipoles with the source point shifted into the complex domain. The purpose of this study is twofold: (i) to develop an understanding of the major characteristics of the multipole fields; and, (ii) to find the parameters that adjust these characteristics in such a manner that they are optimally suited for the scattering problem at hand.

The time harmonic wave equation has a set of independent solutions[18] given by

$$f_{mn}(\mathbf{r}) = h_n^{(2)}(kr) P_n^m(\cos \theta) \exp(jm\varphi) \qquad (1)$$

and known as spherical harmonics, where (r, θ, φ) are the spherical coordinates of the observation point $\mathbf{r}$. In (1), $P_n^m(\cdot)$ denotes the associated Legendre polynomial of order n, m, $h_n^{(2)}(\cdot)$ denotes the spherical Hankel function of the second kind and order n, k is the intrinsic wave number, and the $\exp(jwt)$ time dependence is implicit. These solutions have singularities at the origin of the coordinate system. Thus, the spherical harmonics can represent the acoustic pressure field at an observation point $\mathbf{r}$, produced by a time-harmonic multipole source of order (n,m) located at the source point $\mathbf{r}_0 = 0$.

In order to analytically continue f_{mn} to the case of an arbitrary complex source point $\mathbf{r}_0$, it is necessary to express them as single-valued analytic functions that are well defined for an arbitrary complex source point $\mathbf{r}_0 = (x_0, y_0, z_0) = \mathbf{r}_0' - j\mathbf{r}_0''$, where $\mathbf{r}_0'$ and $\mathbf{r}_0''$ are real vectors. To this end, the following expression,

$$r = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2} \quad \text{Re}\{r\} \geq 0, \qquad (2)$$

which interprets r for an arbitrary $\mathbf{r}_0$, was introduced[13] in the study of the zeroth-order multipole with a complex source point. The expression under the square root in (2) vanishes on a circle of radius $|\mathbf{r}_0''|$ lying in the plane normal to $\mathbf{r}_0''$ and centered at $\mathbf{r}_0'$. The requirement $\text{Re}\{r\} \geq 0$ defines the disc within that circle as a branch cut. It is well known that the radiation characteristics of the zeroth-order multipole approximate those of a Gaussian beam propagating in the direction of $\mathbf{r}_0''$ with the waist located on the branch cut.

As mentioned earlier, Shin and Felsen[19] have proposed a derivation of the higher-order multipole fields based on successive derivatives of the field of zeroth-order multipole along the directions transverse to $\mathbf{r}_0''$. However, this process leads to only a partial set of multipoles, although additional multipoles can be obtained by taking the derivatives along $\mathbf{r}_0''$. In this work, we use an alternative approach, which we believe to be a simpler way to derive the multipole fields. The analytic continuation procedure entails a search for analytic expressions for the Cartesian components of the spherical harmonics (1) in terms of Cartesian coordinates of the source and observation points. The interpretation of θ and φ in (1) as angles in the complex source point case may not be immediately apparent. However, we observe that the trigonometric functions and the exponential factor are easier to continue in their entirety when rewritten as follows:

1320   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Boag *et al.*: Multipole approach to scattering   1320
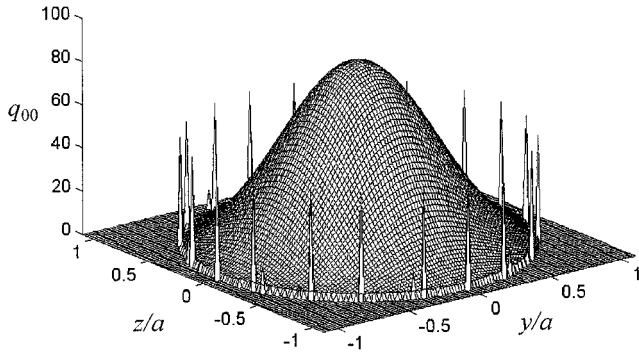
FIG. 1. Distribution of the equivalent double layer source for the f_{00} multipole located at $\mathbf{r}_0 = -ja\hat{\mathbf{x}}$ over a disc of radius a in the yz plane.

FIG. 2. Contour plots of the absolute value of the multipole field $f_{nn}(\mathbf{r} - \mathbf{r}_0)$ computed on a sphere $|\mathbf{r}| = \tilde{r} = 3\lambda$ with $\mathbf{r}_0 = -j\hat{\mathbf{x}}$ for various values of n.

$$\cos \theta = \frac{z - z_0}{r}, \quad \sin \theta = \frac{\rho}{r},$$

$$\cos \varphi = \frac{x - x_0}{\rho}, \quad \sin \varphi = \frac{y - y_0}{\rho} \tag{3}$$

and

$$\exp(jm\varphi) = (\cos \varphi + j \sin \varphi)^m, \tag{4}$$

where

$$\rho = \sqrt{(x - x_0)^2 (y - y_0)^2} \quad \text{Re}\{\rho\} \geq 0. \tag{5}$$

Note that ρ will not appear in the final expressions, and thus the singularity due to the square root in (5) is of no consequence. Using relations (3)–(5), one can obtain the sought-for expressions for Cartesian components of multipole fields in terms of the Cartesian coordinates of the observation and source points. Examination of the fields produced by the multipoles shifted into the complex domain have shown that for $n > 1$ only multipoles with equal indices ($m = n$) produce well-defined beams. Thus, we present here explicit expression only for that case, simplified thanks to the form of associated Legendre polynomials with equal indices, namely $P_n^n(\cos \theta) = (-1)^n (2n-1)!! \sin^n \theta$, where $(2n-1)!!$ denotes the product of all odd positive integers not greater than $2n - 1$. We have

$$f_{nn} = (-1)^n (2n-1)!! h_n^{(2)}(kr) E^n \tag{6}$$

and, in addition,

$$\nabla f_{nn} = (-1)^n (2n-1)!! \left\{ (\hat{\mathbf{x}} + j\hat{\mathbf{y}}) \frac{n}{r} h_n^{(2)}(kr) E^{n-1} \right.$$

$$\left. - k h_{n+1}^{(2)}(kr) E^n (r_x \hat{\mathbf{x}} + r_y \hat{\mathbf{y}} + r_y \hat{\mathbf{z}}) \right\}, \tag{7}$$

where we define $r_x = (x - x_0)/r$, $r_y = (y - y_0)/r$, $r_z = (z - z_0)/r$, and $E = r_x + j r_y$. Equations (6) and (7), with r defined as in (2), are useful for constructing the acoustic pressure and velocity fields of the nnth order multipole located at an arbitrary complex source point. To gain better understanding of the properties of the multipoles located in complex space we first examine their equivalent real space sources. The requirement $\text{Re}\{r\} \geq 0$ in (2) causes r to be discontinuous over the branch cut disc of radius $|\mathbf{r}_0''|$ lying in the plane

normal to $\mathbf{r}_0''$ and centered at $\mathbf{r}_0'$. Consequently, both pressure and normal velocity are discontinuous over that disc. The discontinuity of the fields can be interpreted as single and double layer surface sources distributed in the circular aperture. As will be shown below, this combination of these single and double layer sources can produce unidirectional radiation similar to that in the combined source formulation of the IML.[4] An example of the equivalent double layer source distribution for the f_{11} multipole located at $\mathbf{r}_0 = -ja\hat{\mathbf{x}}$ is illustrated in Fig. 1. Here, one can see that the source is characterized by a smooth distribution over the disc and a singular behavior on the circle $y^2 + z^2 = a^2$.

Under complex shift of the source point, the multipole fields become well-defined beams. The qualitative behavior of the f_{nn} beams in the spatial domain is depicted in Fig. 2 for several values of n. Here, the f_{nn} computed on a sphere with radius $|\mathbf{r}| = \tilde{r} = 3\lambda$ (λ being the wavelength) are presented as functions of θ and φ for $\mathbf{r}_0'' = 1\lambda \hat{\mathbf{x}}$ and $\mathbf{r}_0' = 0$. Examination of Fig. 2 reveals that the complex multipoles produce well-confined beams. However, for higher-order multipoles these beams do not propagate in the $\mathbf{r}_0''$ direction. Spectral analysis of the multipole beams shows that they are also well confined in the spatial frequency domain (see Fig. 3). The beams exhibit a traveling wave behavior in the φ direction, with an increasing spatial frequency for ascending n. The behaviors of these beam fields on the sphere resemble
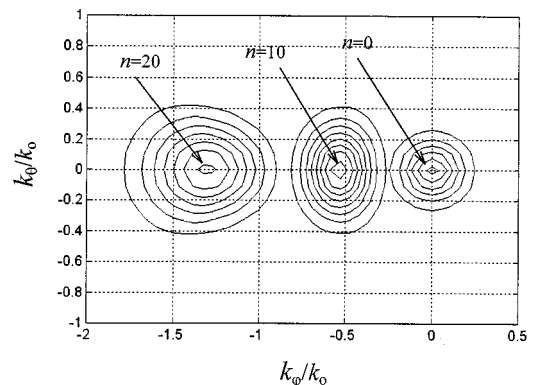
FIG. 3. Contour plots of the absolute value of the 2D discrete Fourier transform of the field distributions presented in Fig. 2.

those of the 2D Gabor expansion basis functions,[17] which comprise a 2D window function multiplied by an exponential factor. Identifying the similarity between the Gabor functions and the multipole fields is of major importance to us, as it provides us with the guidelines for selecting the parameters that play a critical role in our approach to solving the scattering problem. For the multipole fields, the shift of the origin into the complex domain has an effect similar to that of the window function. Unlike in the Gabor expansion, we do not have the freedom to select the window function; however, we can adjust the amount of shift into the complex domain and select multipole orders of the complex multipole expansion. In the remainder of this section, we will describe how to select a set of multipoles that approximate the set of Gabor functions.

It is evident from Fig. 2 that the multipole beams of different orders have different widths. The average beam-width can be adjusted to some predetermined value W by varying the amount of complex shift. We choose to use only the multipoles of orders n_l, $n_l = lK$, $l = 0,1,...,$ where the integer K is determined by the condition $K < 2\pi\tilde{r}/W$. This condition, based on an assumption that the oscillatory behavior of $e^{jn\varphi}$ is preserved, ensures certain oversampling.[16] The multipole series can be truncated by limiting l to the range $0 \leq l \leq L$, where L is the smallest integer satisfying $KL > k\tilde{r}$. The above condition implies that the spatial spectrum of at least one of the multipoles falls outside of the visible range.

By analogy with Ref. 16 we define

$$W_l = \left(\frac{\int_S |f_{n_l n_l}(\mathbf{r})| d^2 s}{\max_{r \in S} |f_{n_l n_l}(\mathbf{r})|} \right)^{1/2} \tag{8}$$

as an effective width of the beam shaped field $f_{n_l n_l}$. Here, the source point is centered at the origin $\mathbf{r}_0' = 0$, and shifted into complex space by $\mathbf{r}_0'' = \mathbf{a}_l = (a_l \cos a_l, a_l \sin a_l, 0)$. The surface S, on which $f_{n_l n_l}(\mathbf{r})$ is evaluated, is a smooth surface that approximates the surface of the scatterer in some average manner. The amount of shift a_l into the complex domain is found for each multipole of order n_l such that the beams are approximately equal in their effective width, i.e., $W_l \approx W$. The direction of the complex shift of each multipole α_l is chosen to ensure that the beam propagates in the x direction. With the choice of these carefully tuned parameters K, L, W, and $\mathbf{a}_l$, the multipole fields comprise a basic set of analytic solutions of the wave equation which, when evaluated on S, simulate a basic set of Gabor functions. A complete set of Gabor-like functions can then be generated by rotation and shift of these functions. This step of constructing the above basis functions serves as a first step toward developing the formulation of the scattering problem, to be discussed in the following section.

II. FORMULATION

For the sake of illustrating the application of the proposed method, we consider an impenetrable body of arbitrary shape, whose geometry, and that of the coordinate system used, are depicted in Fig. 4. The boundary of the body, de-
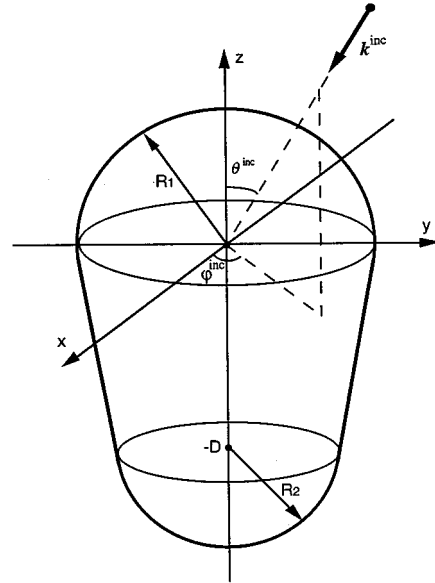
FIG. 4. An example of impenetrable closed body of arbitrary smooth shape illuminated by a plane wave.

noted by S, can be either perfectly hard or soft or can be characterized by an impedance boundary condition. In addition, S is assumed to be piecewise smooth, with a minimum radius of curvature that exceeds some prescribed value $\tilde{r}$. The body is illuminated by a plane wave $P^{\text{inc}} = \exp(-j\mathbf{k}^{\text{inc}} \cdot \mathbf{r})$, where $\mathbf{k}^{\text{inc}}$ denotes the wave vector of the incident field and the harmonic $\exp(j\omega t)$ time dependence is implicit. Our objective is to determine the scattered field P^s.

Following the basic strategy described in Refs. 7–10, we begin by setting up an equivalent problem for the region surrounding the scatterer. Here, we express P^s and the corresponding velocity field $\mathbf{v}^s$ as superpositions of fields of a set of fictitious sources, located in the region originally occupied by the scatterer. Specifically, using the complex multipole sources, we write

$$P^s(\mathbf{r}) = \sum_{i=1}^{M} \sum_{l=0}^{L} \sum_{v=1}^{N(l)} A_{ilv} P_{ilv}(\mathbf{r}) \tag{9}$$

and

$$\mathbf{v}^s(\mathbf{r}) = \sum_{i=1}^{M} \sum_{l=0}^{L} \sum_{v=1}^{N(l)} A_{ilv} \mathbf{v}_{ilv}(\mathbf{r}), \tag{10}$$

where A_{ilv} are constant coefficients to be determined,

$$P_{ilv}(\mathbf{r}) = f_{n_l, n_l}(\underline{\mathbf{R}}_{ilv}(\mathbf{r} - \mathbf{r}_i') + j\mathbf{a}_l), \tag{11}$$

and

$$\mathbf{v}_{ilv}(\mathbf{r}) = -\frac{1}{j\omega\rho} \underline{\mathbf{R}}_{ilv}^T \nabla f_{n_l, n_l}(\underline{\mathbf{R}}_{ilv}(\mathbf{r} - \mathbf{r}_i') + j\mathbf{a}_l). \tag{12}$$

Here, P_{ilv} and $\mathbf{v}_{ilv}$ are, respectively, the pressure and velocity fields of the multipole of order n_l, with $n_l = lK$, located at $\mathbf{r}_{il} = \mathbf{r}_i' - j\underline{\mathbf{R}}_{ilv}^T \mathbf{a}_l$. Also $\mathbf{a}_l$ is the complex location of the multipole of order n_l in its local coordinate system, and matrix $\underline{\mathbf{R}}_{ilv}$ is an orthogonal matrix ($\underline{\mathbf{R}}_{ilv}^T = \underline{\mathbf{R}}_{ilv}^{-1}$) transforming a vector from the global coordinate system to the local one. Note that, unlike in the MMP approach, only the multipoles
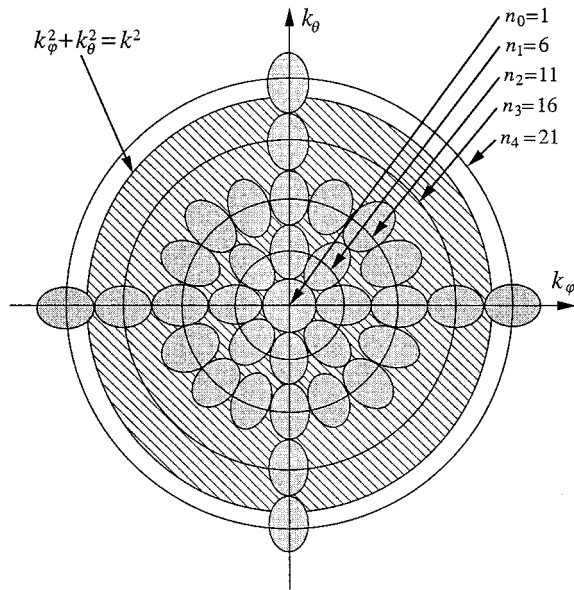
FIG. 5. The local k-space polar grid of an approximate 2-D Gabor expansion based on the multipole beams with $K=5$ and $L=4$. The areas covered by the multipole beams are schematically shown as shaded ellipses. Only few out of $\mathbf{N}(l)$ beams of order n_l are shown for $l=3,4$.

whose orders are integer multiples of K and whose indices are equal, are employed herein. The source locations are chosen by analogy with the Gabor expansion in the following manner: The surface of the scatterer is first subdivided into M patches with linear dimensions not exceeding the predetermined effective width W of the multipole beams. The multipole centers $\mathbf{r}_i'$ are placed at the centers of these patches, and are shifted an equal distance $\widetilde{r}$ toward the interior of the scatterer, namely $\mathbf{r}_i'=\mathbf{r}_i^s-\widetilde{r}\ \hat{\mathbf{n}}_i$, where $\mathbf{r}_i^s$ is an approximate center of the ith patch and $\hat{\mathbf{n}}_i$ is the unit outward normal vector at $\mathbf{r}_i^s$. The value of $\widetilde{r}$ is selected such that it does not exceed the minimum radius of curvature. The $\mathbf{N}(l)$ fields f_{n_l,n_l} of the multipoles of order n_l centered at $\mathbf{r}_i'$ are easily computed in their local coordinate systems. These local coordinate systems are rotated relative to the global one so that their x axes coincide with the normal vector $\hat{\mathbf{n}}_i$. They are also rotated relative to each other around their common x axis. Thus, the matrix $\underline{\mathbf{R}}_{il\upsilon}$ performs first rotation around the z and y axes to align the x axis with $\hat{\mathbf{n}}_i$, and then around the x axis by angle $2\pi\upsilon/\mathbf{N}(l)$. The number of rotations $\mathbf{N}(l)$ is given by $\mathbf{N}(l)=8l$, $l\geq1$, while $\mathbf{N}(0)=1$. The vectors $\mathbf{a}_l$ defining the amount and direction of the shift into the complex domain in the local system are predetermined in a manner described in the previous section.

The completeness of the multipole beams as the sources for the scattered field depends upon whether the functions obtained by sampling these beams on the scatterer boundary form a complete set for functions defined on that boundary. For the case of multipole beams, the completeness stems from their similarity with the Gabor basis set. We anticipate that the similarity between the set of complex multipole beams, selected according to the procedure described in the previous section, and the Gabor basis functions is preserved for an arbitrarily smooth surface. Figure 5 explains how a set of multipole beams originating at $\mathbf{r}_i'$ spans the spatial fre-

quency space local to the ith patch. Like in Fig. 3, k_θ and k_φ denote the local spatial frequencies. The areas in this k space covered by the multipole beams are schematically shown as shaded ellipses. For this example we selected $K=5$ and $L=4$. Due to the large value of $\mathbf{N}(l)$ for $l=3,4$ only a few of these beams are shown. The multipole beams (the ellipses in Fig. 5) are required to cover the spatial bandwidth of the field sampled on the boundary. The latter extends only slightly beyond the visible spectrum shown as a cross hatched circle of radius k in Fig. 5. The criterion for the truncation of the beam series (see Sec. I) has been designed precisely to accommodate this behavior. The highest-order multipoles are selected so that they produce mainly evanescent waves, i.e., in Fig. 5 the centers of the outermost shaded ellipses are located outside the circle of radius k. With this distribution of sources, each patch and, consequently, the whole boundary, are fully spanned by the multipole beams. In view of this, the representation of the scattered field given in (9) and (10) may be viewed as an approximate Gabor expansion evaluated on the scatterer boundary. Note that, in the conventional 2D Gabor expansion employed in aperture radiation modeling[15] and in image processing applications,[17] the spatial and the spatial frequency domains are sampled on Cartesian grids. Here, we find it more convenient to cover the frequency domain on a polar grid by rotating the basic functions as shown in Fig. 5. A similar polar scheme has been suggested in Ref. 17. A log–polar scheme has also been used to construct the 2D ''wavelet'' transform.[22]

It is evident from (9), that the approximate scattered field P^s automatically satisfies both the wave equation and the radiation condition. Thus, assuming a total of N unknown coefficients, it is only necessary to enforce the appropriate boundary condition at N equispaced points on the boundary of the scatterer to derive a matrix equation which can be subsequently solved for the unknown coefficients. Once these are known, the scattered fields and other quantities of interest can be readily obtained.

At this point, it would be useful to make a few observations regarding the method just outlined. First, we note that, thanks to the properties of multipole fields, the generalized impedance matrix involved is anticipated to have a well-defined structure. For a convex scatterer, the matrix can be expected to be essentially banded with enhanced block-diagonal elements. This pattern is attributable to the strong coupling between each group of multipoles having a common origin and the matching points closest to them. On the convex portions of the scatterers, the coupling rapidly diminishes with increasing distance between the multipole origin and the match point. The banded structure of the matrix offers a number of advantages in terms of computation time and storage requirements. Assuming square patch shape, we can estimate the lower bound on the number of unknowns per square wavelength of the surface of the scatterer. Noting that there are $(2L+1)^2$ match points per patch of area W^2, and using the constraints on the multipole parameters specified in the previous section, we can show that this bound is given by $(2L+1)^2/L^2$. This bound approaches the limit of four unknowns per square wavelength for larger L, which is equivalent to the Nyquist limit of two unknowns per linear

wavelength. However, large values of L can only be employed for very smooth scatterers, and even then they tend to spoil the banded structure of the matrix.

III. NUMERICAL RESULTS

The formulation presented in the preceding section has been implemented in a computer program. In this section, we present the computed results for scattering from perfectly soft and hard spherical bodies. Though in our examples we consider only spherical scatterers to simplify the geometrical modeling, solution of the acoustic problem treats them as general 3D scatterers. The accuracy of the proposed method is verified by comparing our numerical results with those derived by using the Mie series. For more complex scatterers, for which analytic solutions are unavailable, an internal validity check is provided by evaluating the error in the satisfaction of the boundary condition on the surface of the scatterer at locations between the match points.

In all of the computations, presented herein, we have employed the same basic set of multipole beams. We would typically begin by optimizing the parameters of the multipoles producing these beams. We would displace the multipole centers by $\tilde{r} = 3\lambda$ from the scatterer surface and choose a beamwidth of $W = 3\lambda$ as a compromise between a reasonable directivity and locality of the interaction. For the surface S we would choose a sphere 4.5λ in radius for the effective area computation as in (8). The center of S would be offset in the negative x-direction by 1.5λ away from the multipole origin. For these choices, we would next optimize the remaining parameters by following the procedure described towards the end of Sec. I. To ensure sufficient oversampling, we would choose $K = 5$ and, consequently, L would be 4. The results obtained by following the above procedure are presented in Table I.

The first example considered was that of a perfectly soft sphere of radius $R = 4.2\lambda$. In accordance with the guidelines provided in the last section, $N = 1670$ multipoles were distributed on a sphere of radius 1.2λ and concentric with the boundary. With a scatterer surface area of 221.7 square wavelength, the average number of unknowns per square wavelength of the surface was 7.5, which, though somewhat higher than the lower bound derived in the previous section, viz., $(2L+1)^2/L^2 = 5.1$ for $L = 4$, is still very low compared to the conventional MM. The structure of the resulting matrix is illustrated in Fig. 6, showing relative magnitude of the elements in the first column plotted versus the row index. The first group of about 61 elements corresponds to the matching points on the patch illuminated by the first multipole. The next six groups of elements are due to the five

TABLE I. Multipole parameters optimized for $W \approx 3\lambda$ on a spherical surface 4.5λ in radius with $\tilde{r} = 3\lambda$, and $K = 5$.

l	n_l	a_l/λ	α_l (deg)
0	0	1.15	0
1	5	0.80	16
2	10	0.60	33
3	15	0.45	56
4	20	0.15	83
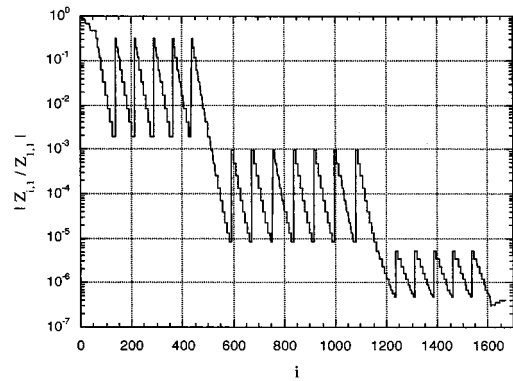
FIG. 6. Normalized absolute value of the elements in the first column of the generalized impedance matrix versus row index for the case of the perfectly soft sphere of radius $R = 4.2\lambda$.

nearest neighbor patches. The magnitude of the elements in all other groups corresponding to further patches is relatively low (order of 10^{-3} and lower). Consequently, these elements can be set to zero with a negligible loss of accuracy. The bandwidth of nonzero elements in the matrix becomes approximately 500 when this is done. In view of this, we estimate that for a convex scatterer the storage requirement will only be on the order of $500N$, and, consequently, will result in a reduction of computation time. The root mean square (rms) boundary condition error for this case was found to be below 0.6%. A comparison was also made between the scattering cross section results computed by our method and those obtained by using the exact modal expansion. These results are shown in Fig. 7 for the case of plane wave incident at $\theta^{\text{inc}} = 0°$. Complete agreement between the two serves as additional evidence of the accuracy of the proposed method.

The second problem considered was that of a perfectly hard sphere of radius $R = 4.2\lambda$. For this case, we again used a total of $N = 1670$ unknowns. The rms boundary condition error was only about 1.3%, and this provided an internal check on the accuracy of the method. In Fig. 8, the accuracy of the scattering cross section computed by our method is exemplified again by comparing it with one derived by using the Mie series. From the two examples one can see that the method described in this paper provides highly accurate results for scattering by moderately large smooth bodies. These
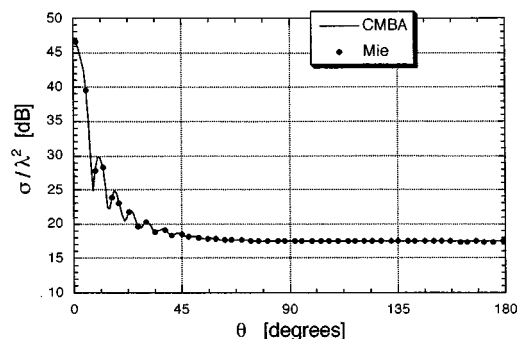
FIG. 7. Scattering cross section of the perfectly soft sphere of radius $R = 4.2\lambda$ illuminated by a plane wave incident at $\theta^{\text{inc}} = 0°$.
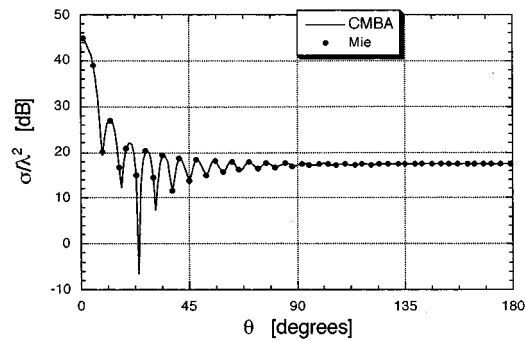
FIG. 8. Scattering cross section of the perfectly hard sphere of radius $R = 4.2\lambda$ illuminated by a plane wave incident at $\theta^{inc} = 0°$.

are the cases where the use of conventional BE/MM can be computationally very costly, while asymptotic techniques aren't accurate enough.

IV. DISCUSSION

A novel multipole-beam approach for a full-wave, 3D analysis of acoustic scattering from arbitrary shaped bodies has been proposed in this paper. In this method, the scattered field is expanded in a series of beams produced by multipole sources located in the complex space. Advantages of the proposed approach over the existing methods, in terms of the number of unknowns, have been demonstrated for a number of examples. In its present form, the method is most suitable for the analysis of large smooth bodies, for which it can provide substantial savings of computer memory and time over the conventional boundary element/moment method. A hybrid approach to extending the present method to a more general class of scatterers is currently under investigation and the initial results appear to be quite promising.[23]

[1] T. W. Wu, "A direct boundary element method for acoustic radiation and scattering from mixed regular and thin bodies," J. Acoust. Soc. Am. **97**, 84–91 (1995).

[2] R. F. Harrington, *Field Computation by Moment Methods* (Macmillan, New York, 1968).

[3] F. X. Canning, "The impedance matrix localization (IML) method for moment-method calculations," IEEE Antennas Propag. Mag. **32**, 18–30 (1990).

[4] F. X. Canning, "A new combined field integral equation for impedance matrix localization (IML)," in *Proceedings of the IEEE–AP/S International Symposium* (Ann Arbor, Michigan, July 1993).

[5] E. M. Rosen, F. X. Canning, and L. S. Couchman, "A sparse integral equation method for acoustic scattering," J. Acoust. Soc. Am. **98**, 599–610 (1995).

[6] D. Gabor, "Theory of communication," J. Am. Inst. Electr. Eng. **93**, 429–457 (1946).

[7] C. Hafner, *The Generalized Multipole Technique for Computational Electromagnetics* (Artech House, Norwood, MA, 1990).

[8] M. G. Imhof, "Multiple multipole expansions for acoustic scattering," J. Acoust. Soc. Am. **97**, 754–763 (1995).

[9] A. Boag, Y. Leviatan, and A. Boag, "Analysis of acoustic scattering from fluid cylinder using a multifilament source model," J. Acoust. Soc. Am. **83**, 1–8 (1988).

[10] A. Boag, Y. Leviatan, and A. Boag, "Analysis of acoustic scattering from fluid bodies using a multipoint source model," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **36**, 119–128 (1989).

[11] R. J. Pogorzelski, "Improved computational efficiency via near-field localization," *Proceedings of the URSI Radio Science Meeting* (Chicago, July 1992).

[12] E. Erez and Y. Leviatan, "Analysis of scattering from structures containing a variety of length-scales using a source-model technique," J. Acoust. Soc. Am. **93**, 3027–3031 (1993).

[13] G. A. Deschamps, "Gaussian beams as a bundle of complex rays," Electron. Lett. **7**, 684–685 (1971).

[14] P. D. Einziger and S. Raz, "Beam-series representation and the parabolic approximation: The frequency domain," J. Opt. Soc. Am. A **5**, 1883–1892 (1988).

[15] J. M. Klosner, L. B. Felsen, I. T. Lu, and H. Grossfeld, "Three-dimensional source field modeling by self-consistent Gaussian beam superposition," J. Acoust. Soc. Am. **91**, 1809–1822 (1992).

[16] J. Wexler and S. Raz, "Discrete Gabor expansions," Signal Process. **21**, 207–220 (1990).

[17] Y. Y. Zeevi and M. Porat, "The generalized Gabor scheme of image representation in biological and machine vision," IEEE Trans. Pattern. Anal. Mach. Intell. **10**, 452–468 (1988).

[18] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941).

[19] S. Y. Shin and L. B. Felsen, "Gaussian beam modes by multipoles with complex source points," J. Opt. Soc. Am. **67**, 699–700 (1977).

[20] A. Boag and R. Mittra, "Complex multipole beam approach to 3D electromagnetic scattering problems," J. Opt. Soc. Am. A **11**, 1505–1512 (1994).

[21] A. Boag and R. Mittra, "Complex multipole beam approach to electromagnetic scattering problems," IEEE Trans. Antennas Propag. **42**, 366–372 (1994).

[22] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," IEEE Trans. Acoust. Speech Signal Process. **36**, 1169–1179 (1988).

[23] A. Boag, E. Michielssen, and R. Mittra, "Hybrid mutipole-beam approach to electromagnetic scattering problems," J. Appl. Comp. Electromag. Soc. **9**(3), 7–17 (1994).

Seat dip effect: The phenomena and the mechanism

D. Takahashi

Department of Architectural Engineering, Fukui University, Bunkyo 3-9-1, Fukui 910, Japan

The phenomenon of excess attenuation due to audience seating, which is called ''seat dip effect (SDE),'' have been discussed in the case of a rather short-time interval including a direct sound and early arriving sound. However, it may be expected that the effect is diminished as the time interval increases, owing to the supplement of other arriving sounds. The first aim of this investigation is to clarify the effect of time interval on the SDE as well as the effect of other arriving sounds, by means of $\frac{1}{10}$-scale model experiments. For the mechanism of the phenomenon, a clear explanation has not yet been given. The second aim is to clarify the fundamental mechanism of the SDE. For this aim, a wave-scattering theory for a periodic surface and a simple wave reflection theory for a layered boundary surface are applied, and discussed in comparison with the experimental results. From this investigation, it became clear that the mechanism of the SDE can be easily explained by using these simple analytical models. © *1997 Acoustical Society of America.* [S0001-4966(97)00509-2]

PACS numbers: 43.20.Fn, 43.55.Br, 43.55.Ev [JDQ]

INTRODUCTION

When sound waves are passing over audience seats, the sound energy of low frequencies is attenuated, especially at the early stage of the arriving sounds. The phenomena, which are known collectively as ''seat dip effect (SDE),'' were originally reported in two papers,[1,2] and have been investigated both theoretically and experimentally. Recently, an overall review of these studies was presented by Ishida.[3] However, it does not seem that the actual state of the phenomena and the mechanism of the SDE have been fully clarified. More complete understanding is clearly required for the design of auditorium acoustics.

In the past studies, the phenomena have been discussed in the case of a limited time-interval that includes a direct sound and early arriving sounds (delayed at most 50 ms). In actual hearing conditions in a closed space, the degree of attenuation may decrease due to the supplement of later arriving sounds. Regarding this point, Barron[4] states that the bass sound attenuation at an early stage may affect the total energy of low frequencies and has serious subjective effects. Then he recommends a rise in bass frequency reverberation to compensate for the attenuation effect. Apart from this extending to the reverberant region, it seems that the effect of time interval taken for evaluating the SDE plays an important role for evaluating the subjective effect. On this point, Bradley[5] discussed the importance of early/late sound ratios concerning the strength G, in which a time interval of 40 ms was used to evaluate the SDE and to find possible methods for minimizing the effect.

The present paper is divided into two parts. The purpose of the first part is to clarify the effect of time interval on the SDE by means of scale-model experiments, and to discuss the SDE variation with the difference in source and receiver position at rather long time intervals. This model of $\frac{1}{10}$-scale of a small hall with 228 seats has removable surfaces (side walls, back wall, and ceiling) and enables investigation of the contribution of each reflected sound separately.

The latter part focuses on the mechanism of the SDE.

The SDE is characterized by many phenomena, including, for example, the shift of primary dip frequency with the vertical angle of incidence, the effect of the horizontal angle of incidence, the effect of underpass transmission, and so on. For these phenomena, however, reported results do not necessarily have the same tendency. However, consensus is substantial for some features of the SDE:

(1) The attenuation occurs at low frequencies with wide band centered about 100–250 Hz.
(2) The effect depends strongly on the vertical angle of incidence as well as the receiver height: as the angle approaches near grazing, the effect of primary dip expands in a frequency region, which may seriously affect our auditory sense.
(3) The effect appears in the cases both with and without an audience.

For clarifying the fundamental mechanism of the SDE, it is necessary to present a clear explanation for at least these three items. Why does the excess attenuation occur? On this point, it may be assumed that the answer is the interference between a direct wave and reflected waves from the audience seating. This can be supported from the clue that one-quarter wavelength corresponds to the seat depth, as was pointed out by Shultz and Watters,[1] Cremer and Müller,[6] and Bradley.[5] The phase difference of 180° can cause the cancellation effect. Then, why do such reflected waves that cause the SDE phenomena occur? Regarding this point, no clear explanation has yet been given.

As for the theoretical investigation of the SDE mechanism, Ando[7] analyzed the problem of wave scattering from a periodical uneven surface, which is a model of seat rows. The analysis for more realistic conditions was presented by Kawai.[8] In this analytical model, the effect of underpass as well as the effect of dimension of the seating area can be evaluated owing to the use of a rigorous integral equation method. These two analyses provide rather detailed models
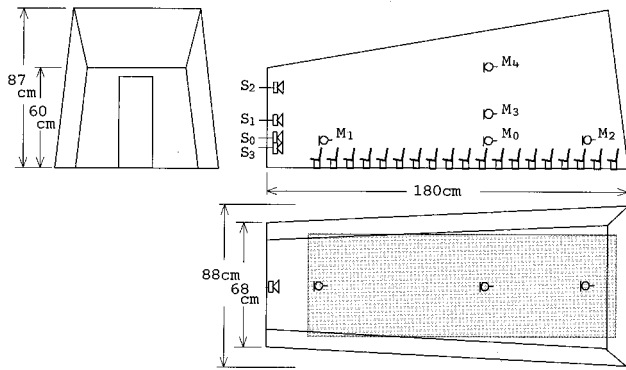
FIG. 1. Configuration of $\frac{1}{10}$-scale model of a small hall.

FIG. 2. Block diagram of the measurement.

of seat rows, but do not provide any clear explanation for the SDE mechanism.

In this investigation, an explicit analytical form for explaining the SDE mechanism, is derived from both a wave-scattering theory for a simple model of seat rows and a simple wave reflection theory for a layered boundary surface. The final result of this analysis, together with the experimental verification, clarifies the fundamental mechanism of the SDE.

I. PHENOMENA OF THE SDE IN AN ARTIFICIAL SPACE

The effect of time interval for evaluating the SDE and the effect of supplemental waves reflected from surroundings were investigated by using a 1/10-scale model of a small hall. The details of the model are shown in Fig. 1. The floor was made of 24-mm plywood panel and the walls and ceiling were 12-mm plywood panels. These panels, which compose an almost perfectly closed space, can be removed separately. The seat of 90-mm height, including a seat back of 50-mm height, was made of polystyrene board. The arrangement of the seats was 12 columns×19 rows with row-to-row spacing of 85 mm. All of these were set in an anechoic chamber. The experimental method is shown schematically in Fig. 2. The sound pressure (SPL) level L normalized by the direct sound can be obtained by performing the measurements with and without the seating area, which is given by

$$L = 10 \log_{10} \left| \frac{H_d + H_r}{H_d} \right|^2 = 10 \log_{10} \left| \frac{\Phi_{12}/\Phi_{11}}{\Phi_{12}^0/\Phi_{11}^0} \right|^2, \qquad (1)$$

where H_d and H_r are the frequency response functions of the direct pass and the reflected pass, respectively. A broadband noise of 500–12.5 kHz, which was fed by a white noise generator with high-pass and low-pass filters, was used as a source. Two signals from noise generator and the microphone output were digitized and recorded, and subsequently processed collectively by a personal computer. For the process of transforming (FFT) the correlation functions, the time interval for evaluating the attenuation effect was adjusted by setting the cutoff time T_c of a window function [Hamming window: $0.54 + 0.46 \cos(\pi t/T_c)$ for $t \leq T_c$ and 0 for $t > T_c$]. As is customarily defined, the power spectrum Φ_{11} and cross spectrum Φ_{12} are the Fourier transform of the auto-

correlation and cross-correlation function, respectively.
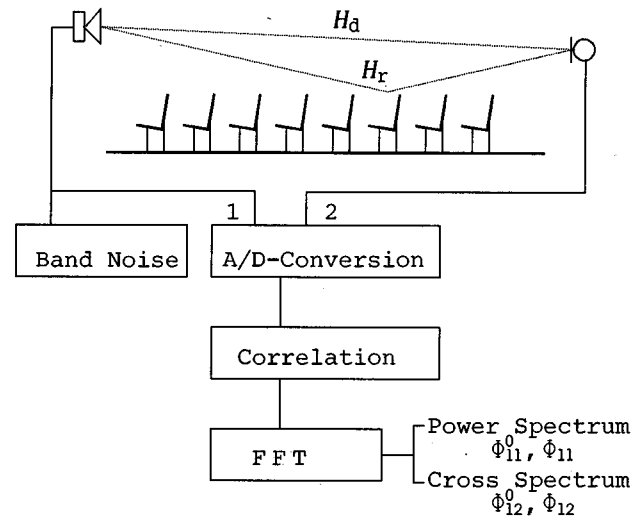
The SDE characteristics at a receiving point M_0 with a parameter of cutoff time of windowing are shown in Fig. 3, in which the source position is S_0. To clarify the effect of the seat arrangement, the results are represented by SPL normalized by the case with the seats removed (not a direct sound, as can be seen in many past studies). Moreover, to remove the additional effect, which is the difference in total energy with and without the seating, the seats were piled up along the side wall in the measurement without seating. As shown in Fig. 3, the degree of attenuation decreases with increased time interval. From the results, it can be seen that the SDE is more serious at an early stage of arriving sound. However, the results of the time interval up to 32 ms (320 ms in full scale), which gives almost an asymptotic state, show a clear dip. The effect on our auditory sense will be a subject of some interest.
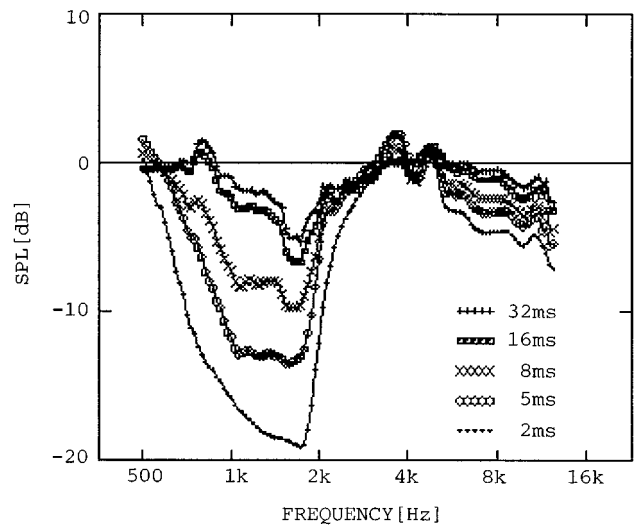
FIG. 3. The SDE characteristics with a parameter of the cutoff time of windowing.
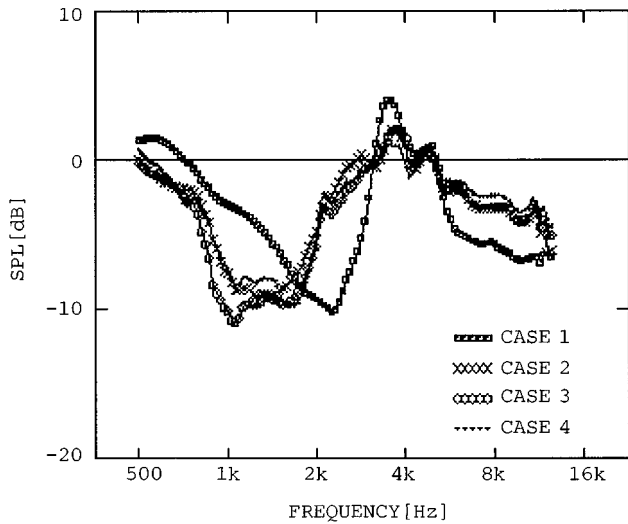
FIG. 4. Effect of reflections from side walls (case 2), side walls+ceiling (case 3), and side walls+ceiling+back wall (case 4). Case 1 represents the floor alone, and each case includes the effect of the front wall.

Figure 4 shows the effect of supplemental waves reflected from other parts. The cutoff time of windowing was fixed at 8 ms (80 ms in full scale). The remarkable differences shown in Fig. 3 diminish in these results. It seems that the differences may be caused by removing all waves outside the interval in the case of Fig. 3. The late arriving waves are considered to be generated from multiple reflections by each composition of this model, which may weaken the attenuation effect. In any case, there are evident excess attenuation phenomena, with dips of about -10 dB in early arriving sound up to 80-ms delay—a range of great subjective importance. From these results in comparison with that shown in Fig. 3, the following suggestions regarding the SDE mechanism may be given: (i) Reflections immediately after the direct sound have a strong cancellation effect, which increases with decreased interval. (ii) The change in horizontal angle of incidence does not have serious effect on the attenuation. This implies fundamentally less influence of the seat backs. These data will be reflected in investigating into the mechanism of the SDE in the next section.

It has been pointed out that the SDE depends strongly on both the source height and the receiver height. This finding for the effect of receiver height may correspond to the subjective difference between the main floor and the front seats on balconies.[2] To make clear these trends, the experiments focusing on the effect of change in source and receiver position were carried out at time interval of 8 ms (80 ms in full scale). It can be seen, from the results shown in Fig. 5, that the SDE is remarkably affected by the change in the vertical angle of incidence, and the effect decreases as the receiver height increases. These results show almost the same tendency as discussed in the past studies using a rather short time interval.[1–3,5]

II. FUNDAMENTAL MECHANISM OF THE SDE

A. Analytical model of seat rows

A simplified model of seat rows, which is composed of parallel rows of vertical slats on a surface with some finite
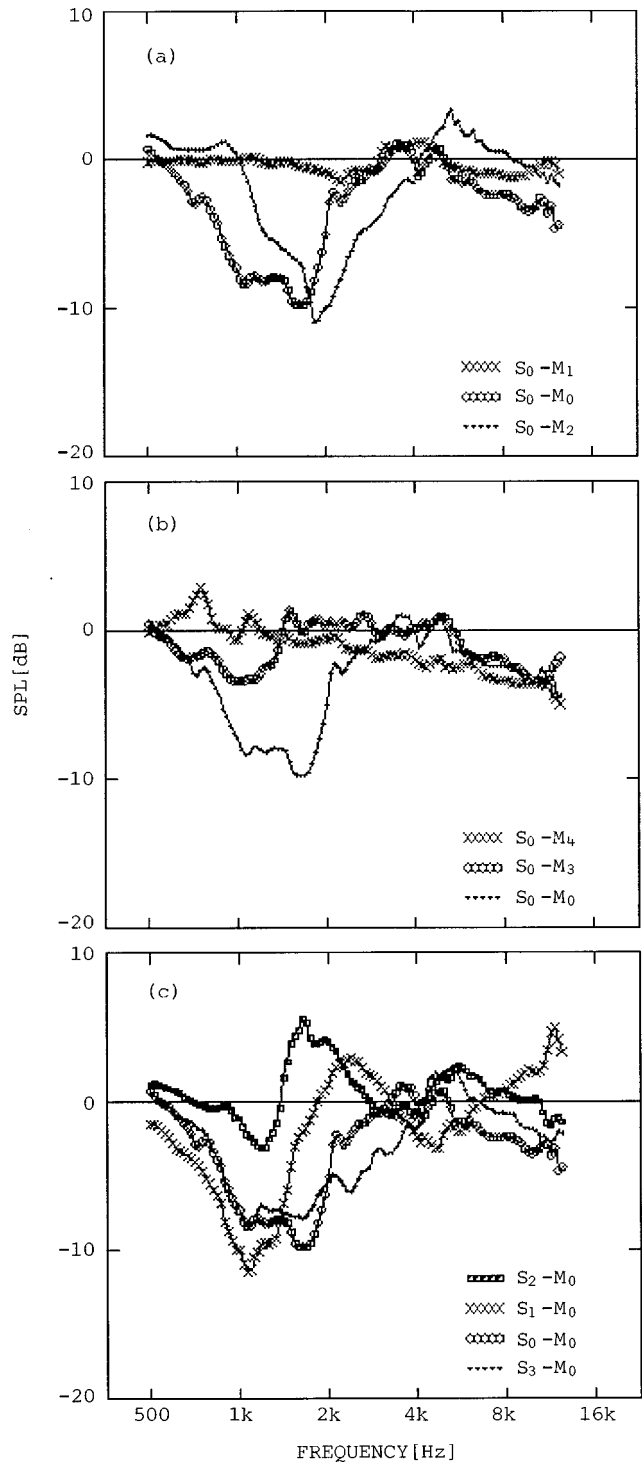
FIG. 5. Effect of source and receiver position: (a) horizontal variation of the receiver, (b) vertical variation of the receiver, and (c) vertical variation of the source.

impedance, is shown in Fig. 6. This configuration of seat rows model was used by Sessler and West[2] and Bradley[5] for measuring the seat dip effect in a scale model. In the model of Sessler and West, vertical wood panels with felt covering were set on a plywood panel, and in Bradley's model, all were made from plywood.

In this section, rigid vertical slats of width w are set on a surface of finite impedance expressed as a specific admit-

1328 J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

D. Takahashi: Seat dip effect 1328
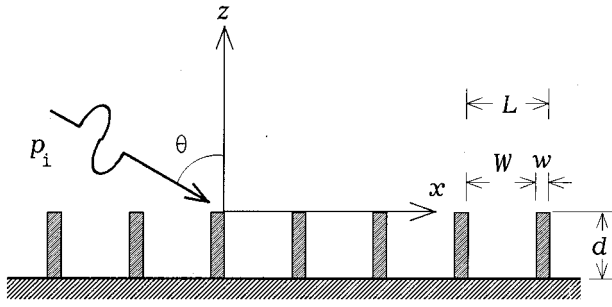
FIG. 6. Geometry of a wave-scattering model for seat rows.

FIG. 7. Effect of vertical-slat width on the attenuation characteristics.

tance A_B. This model may be considered as an extension of Bradley's model. Consider a scattered sound field caused by this structure under an incident plane wave at angle θ, which has the form of sound pressure as $p_i = e^{i(\alpha_0 x - \gamma_0 z)}$ with $\alpha_0 = k \sin \theta$, $\gamma_0 = k \cos \theta$, and $k = \omega/c$ as the wave number, ω and c as the angular frequency and the speed of sound, respectively. The time factor $e^{-i\omega t}$ is suppressed throughout. The boundary surface (plane $z = 0$) is here regarded as a periodically arranged flat surface with two parts of different specific admittance A_1 and A_2 corresponding to the top surfaces of slots and slats, respectively.

The fundamental analysis for finding the solution of the scattered field has been developed by Rayleigh.[9] The scattered wave p_s, respecting the surface periodicity, can be expressed with unknown coefficients Ψ_n as

$$p_s = \sum_{n=-\infty}^{\infty} \Psi_n e^{i(\alpha_n x + \gamma_n z)}, \tag{2}$$

where $\alpha_n = \alpha_0 + 2n\pi/L$ and $\gamma_n = [k^2 - \alpha_n^2]^{1/2}$ with $\mathrm{Re}\{\gamma_n\} \geq 0$ and $\mathrm{Im}\{\gamma_n\} \geq 0$. The boundary condition, by assuming the acoustic characteristics of the surface (plane $z = 0$) to be represented by its specific acoustic admittance A, is

$$\frac{\partial}{\partial z}[p_i + p_s] + iAk[p_i + p_s] = 0 \quad \text{on } z = 0. \tag{3}$$

Generally, the admittance A depends on both the angle of incidence θ and the variable x. In the present case that is given by

$$A = \begin{cases} A_1 : \text{entrance admittance of a slot,} \\ A_2 : \text{surface admittance of a slat } (= 0). \end{cases} \tag{4}$$

Substituting Eq. (2) into Eq. (3) yields

$$\sum_{n=-\infty}^{\infty} \Psi_n(\gamma_n + Ak)e^{i2\pi nx/L} = \gamma_0 - Ak. \tag{5}$$

Transforming this equation using Fourier expansions and rearranging the results with an infinite set of equations with respect to m, one obtains

$$\sum_{n=-\infty}^{\infty} G_{mn}\Psi_n = H_m, \tag{6}$$

where

$$G_{mn} = \frac{1}{L}\int_0^L (\gamma_n + Ak)e^{-i[2(m-n)\pi x/L]}\,dx, \tag{7}$$

$$H_m = \frac{1}{L}\int_0^L (\gamma_0 - AK)e^{-i[2m\pi x/L]}\,dx, \tag{8}$$

and the integration of this type can be done easily in the explicit closed form. If the width of a slot does not exceed half a wavelength, the modal behavior in the x direction can be disregarded.[9] Then the entrance admittance of a slot A_1 can be expressed, for wave propagation in z direction, as

$$A_1 = \frac{A_B - i \tan kd}{1 - iA_B \tan kd}. \tag{9}$$

Numerical examples for the model of $d = 0.4$ m and $W = 0.85$ m are illustrated in Fig. 7, in which the sound pressure level at a point ($x = 0$, $z = 0.2$ m) relative to the direct sound is plotted with two parameters of the incident angle θ and the slat width w. The slot bottom is an absorptive material of thickness 0.1 m and flow resistivity 1×10^4 (mks−rayl/m). By preliminary investigation, it is found that the response is almost independent of the horizontal position x. Though the results fully show the phenomena of SDE, it should be noted that the response remains almost unchanged when $w = 0$. In evaluating the Fourier-type integrals [Eqs. (7) and (8)], when the condition $w = 0$ is introduced, the integrant becomes constant. This means the dominance of 0-order in the Fourier integral. Then the solution of the simultaneous equations of Eq. (6) can be obtained approximately as $\Psi_n \cong (\gamma_0 - Ak)/(\gamma_0 + Ak)$. Consequently, the total sound field can be expressed as

$$p_i + p_s \cong e^{ik \sin \theta \cdot x}\left[e^{-ik \cos \theta \cdot z} + \frac{\cos \theta - A_1}{\cos \theta + A_1}e^{ik \cos \theta \cdot z}\right]. \tag{10}$$

It is also found that the surface represented by Eq. (9) is so-called "locally reacting."
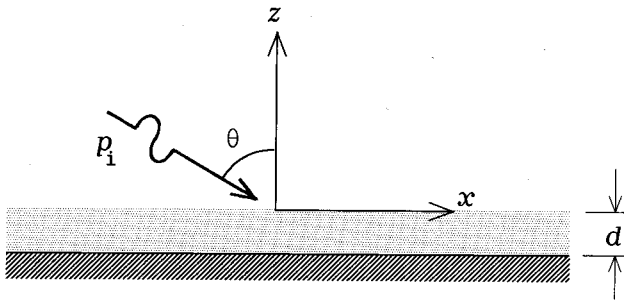
FIG. 8. Geometry of a wave reflection model.

B. Fundamental model of excess attenuation

Consider a sound field reflected by a structure composed of an absorptive layer backed by a rigid floor, under an incident plane wave at angle θ as shown in Fig. 8. The reflected sound pressure p_r is given by

$$p_r = P_r e^{i(\alpha_0 x + \gamma_0 z)}, \tag{11}$$

with an unknown amplitude P_r. If the layer is treated as a homogeneous and isotropic medium of wave propagation, then the sound pressure p and the particle velocity v_z in the z direction inside the layer can be expressed as

$$p = (P_+ e^{-qz} + P_- e^{qz}) e^{rx}, \tag{12}$$

$$v_z = \frac{q}{\gamma Z} (P_+ e^{-qz} - P_- e^{qz}) e^{rx}, \tag{13}$$

where $r = ik \sin \theta$, $q = \gamma \sqrt{1 + (k \sin \theta / \gamma)^2}$ with $\mathrm{Re}\{q\} \geq 0$, $\mathrm{Im}\{q\} \leq 0$, and $P_\pm$ denotes the amplitude associated with the propagating waves toward the $\pm z$ direction, and are unknown quantities. In these expressions γ and Z are the propagation constant and the characteristic impedance of the layer, respectively, which can be calculated by using the equation of Delany and Bazley[10] with the data of flow resistivity. The boundary conditions at both sides of the layer are as follows:

$$v_z\big|_{z=0} = \frac{1}{i\rho\omega} \frac{\partial(p_i + p_r)}{\partial z}\bigg|_{z=0}, \tag{14}$$

$$v_z\big|_{z=-d} = 0, \tag{15}$$

$$[p_i + p_r]_{z=0} = p\big|_{z=0}, \tag{16}$$

with ρ as the air density. Substitution of Eqs. (11)–(13) into these conditions yields the reflected sound pressure p at an arbitrary point (x,z), which is

$$p(x,z) = e^{ik \sin \theta \cdot x} \left[e^{-ik \cos \theta \cdot z} + \frac{\cos \theta - A}{\cos \theta + A} e^{ik \cos \theta \cdot z} \right], \tag{17}$$

where

$$A = \frac{\rho c q}{\gamma Z} \tanh (qd). \tag{18}$$

This quantity is the specific admittance of the boundary surface, and depends on the angle of incidence θ.

The expression of Eq. (17) is identical to Eq. (10), except for the difference in a description of the surface admit-

tance. This is free from the assumption of local reaction, and is so called ''extended reaction.'' The terminology of ''local'' and ''extended'' reaction is used here as is customarily defined (see, for example, Ref. 11). The fundamental mechanism of the excess attenuation can be easily explained by using Eq. (17) or Eq. (10). Let $A = \cos \theta$, for example, which means that the layer is an air of infinite thickness. Then there is only a direct sound. In the case of $A = 0$, the sound pressure at any point can be determined as an ordinary interference phenomenon between a direct wave and a wave reflected from an infinite rigid plane. If A has any other value, a problem occurs: i.e., when the angle θ approaches 90°, whatever the value A may be, the sound pressure p at an arbitrary point nearly equals 0. In the case of the angle near 90°, it can be seen that the effect of attenuation decreases as the receiver height increases. The existence of this type of reflected waves having a cancellation effect of the direct sound has been discussed by Cremer and Müller,[6] Ishida[3,12] and Sakurai et al.[13] In these investigations, however, any theoretical explanation for the correspondence to the seat rows is not given, and the locally reacting assumption is used as the basis, as pointed out by Shultz and Watters.[1] It should be noted that the reflections having this cancellation effect occur in reality without the assumption of the local reaction of the material.

Numerical examples of Eq. (17) are shown in Fig. 9 (curves denoted by ''EXTENDED''), in comparison with the experimental results, which are represented in relative SPL normalized by the direct sound. The absorptive layer is glass wool (80 kg/m³) 20 mm thick and the receiver height is 5 cm above the material, as shown in Fig. 10(a). The curves denoted by ''LOCAL,'' which was calculated using Eqs. (9) and (10), will be discussed in Sec. II C. It is seen from these figures that both theory and experiment show clear dips due to the cancellation effect with dependence on the angle of incidence. However, the theory does not show good agreement with the experiment at low frequencies below 1 kHz. It would be expected that a discrepancy would be caused by the difference in the incident wave front, which is a plane wave in the theory and a spherical wave in the experiment. To confirm this expectation, the sound field reflected from the same locally reacting surface as used in Fig. 9 was calculated by using an approximate solution for a spherical wave incidence.[14,15] The results are shown in Fig. 11, in which the behavior of experimental results at low frequencies seems to be well expressed by this model. The general characteristics, however, are the same as in the case of plane wave incidence. The subsequent analysis used the model for plane wave incidence, because of attractive simplicity with clear perspective.

As can be seen in Eq. (17) and the calculated results of Fig. 9, if the reflecting boundary is not perfectly rigid, the phenomenon of excess attenuation can always occur. To confirm this finding, some experiments with similar conditions were carried out for several different materials. The results are shown in Fig. 12, where the materials used in the experiments are as follows:

(i)   RIGID: nonabsorptive with plywood surface,
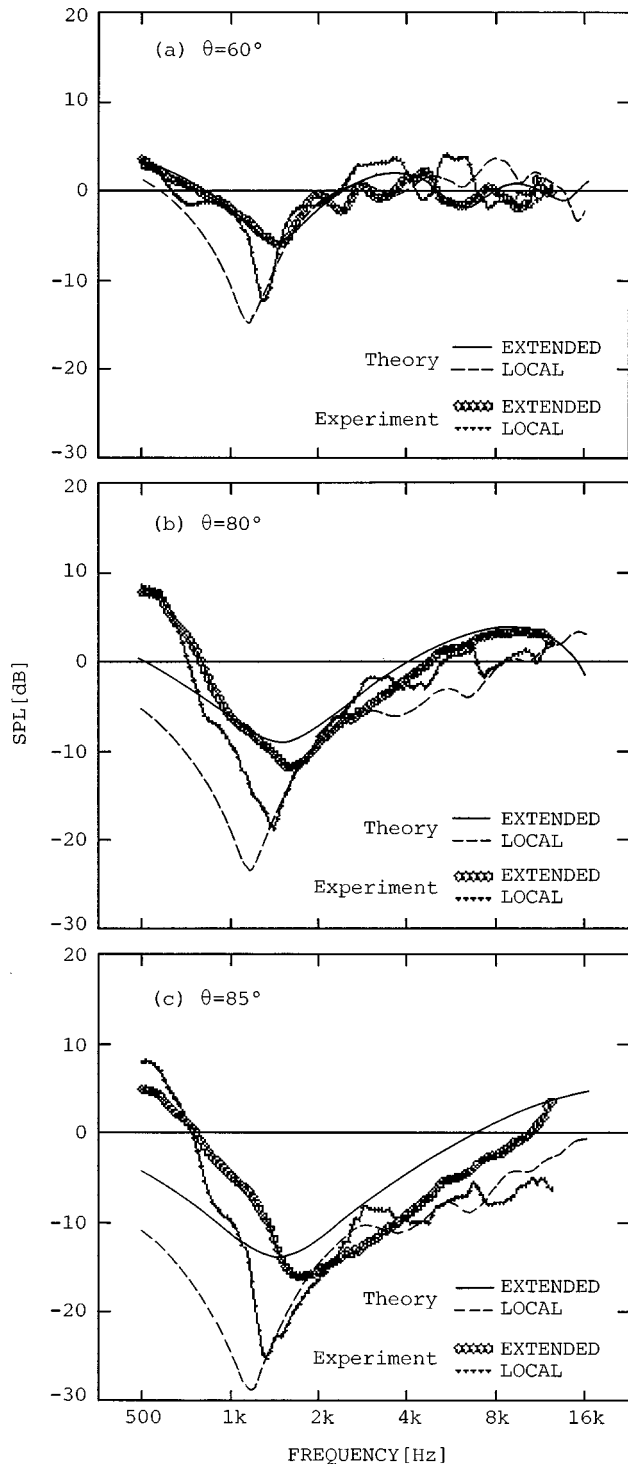
FIG. 9. Attenuation characteristics caused by a sound absorptive layer (GW80-20) with a parameter of incident angle θ: (a) $\theta=60°$, (b) $\theta=80°$, and (c) $\theta=85°$.

(ii)   FR-5: foam rubber 5 mm thick,
(iii)   GW80-12: glass wool ($80 \, \text{kg/m}^3$) 12 mm thick,
(iv)   GW80-20: glass wool ($80 \, \text{kg/m}^3$) 20 mm thick,
(v)   GW32-50: glass wool ($32 \, \text{kg/m}^3$) 50 mm thick.

C. The effect of seat backs

It has been shown in the previous section that, even without the complex unevenness of audience seating, the
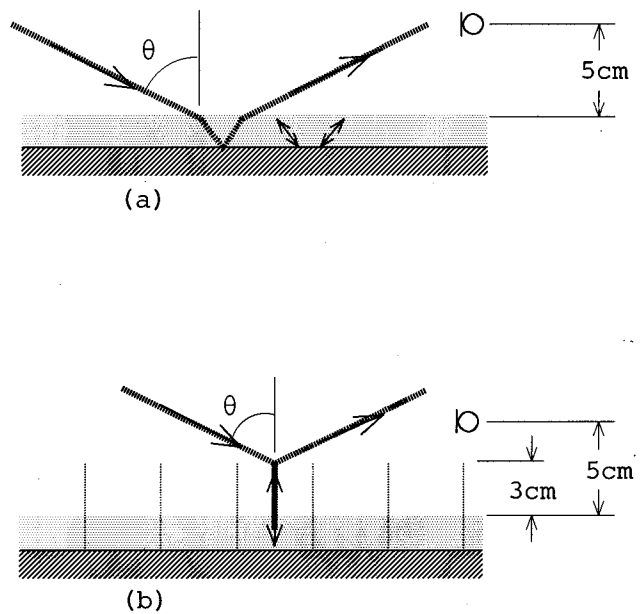
FIG. 10. Geometry of (a) a surface of extended reaction, and (b) a locally reacting surface.

phenomena of excess attenuation occur in reality. However, as can be seen in Fig. 12, frequencies at the maximum dip strongly depend on the difference in material at the slot-bottom. This behavior seems to deviate from the substantial features of the SDE raised in the Introduction, and may be caused by a lack of vertical slats assumed as a simple model of seat backs.

In this section the effect of seat backs on the SDE is investigated both theoretically and experimentally. If the wavelength is longer than twice the row-to-row spacing of the seats, the air-particle movement of the incident wave at any angle of incidence is mainly restricted in the vertical direction. This causes an invisible even surface of impedance discontinuities at the seat tops, a so-called "locally reacting surface." The effect of seat backs can then be evaluated by treating the medium below this surface as a layered material of local reaction, which is represented by Eq. (10) with Eq. (9). The results of this evaluation for the case of Fig. 10(b) are shown in Fig. 9 (curves denoted by "LOCAL") together with the case of extended reaction, in comparison with the experimental data, which were obtained by dividing the material at 5-cm intervals using 5-mm-thick polystyrene boards. The difference between "extended model" and "local model" is schematically shown in Fig. 10. It can be seen from the results in Fig. 9 that the presence of seat backs has the effect of making the dip more serious. Moreover, in comparison with Fig. 12, the seat backs may also have an effect of selective attenuation inherent in the SDE. To confirm these findings, calculations for various bottom materials in the full scale model were carried out with the comparison between "extended" and "local" reaction.

Some examples of the results are shown in Fig. 13, in which the material of the bottom is an absorptive layer 0.1 m thick with a parameter of flow resistivity R (mks-rayl/m),
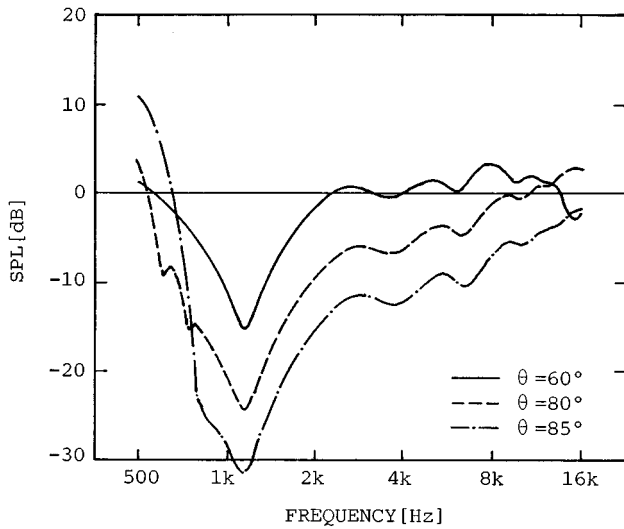
FIG. 11. Attenuation characteristics for a spherical wave incidence.

and the receiving height is 0.6 m from the bottom surface. In the "local" case, the height of the slats (seat backs) is 0.4 m. From this example and other extensive calculations (not shown here), it can be seen that the seat backs have an amplification effect of the excess attenuation and have an effect of characterizing the frequency of primary dip, where the receiver height from the bottom is about one-quarter wavelength. The latter feature means that the phenomena of seat dip attenuation are affected considerably by the seat surface (laps of human bodies for occupied case) rather than the floor surface.

III. POSSIBILITY OF CONTROLLING THE SDE

As was discussed in the previous section, except in the case of a perfectly rigid plane, when the reflecting surface has any acoustic properties the phenomena of the seat-dip attenuation always occur for angles near grazing. From this fact, the SDE occurs even if perfectly rigid seats are arranged on the perfectly rigid floor. In this case, a surface with impedance discontinuities is formed at the seat tops owing to the vertically restrained air-particle movement.

To reduce the attenuation effect, highly absorptive treatment of the floor was proposed.[7] However, this treatment seems not to be an effective measure, because of a strong dependence of the SDE on the seat surface rather than the floor surface, as discussed in Sec. II C. For more details on this point, the effect of absorption of the seat surface was examined by assuming absorption-changeable seats or laps. The results calculated from Eqs. (9) and (10) are shown in Fig. 14(a). The normal-incidence absorption characteristics of the material are shown in Fig. 14(b) for reference. Although it can be seen that peaks and dips are suppressed as the absorption increases in the corresponding frequency range, an attenuation of about −20 dB in a rather wide frequency range is observed in spite of relatively high absorptive treatment, with absorption coefficient of about 0.8–1.0 at frequencies near the maximum dip. This fact means that
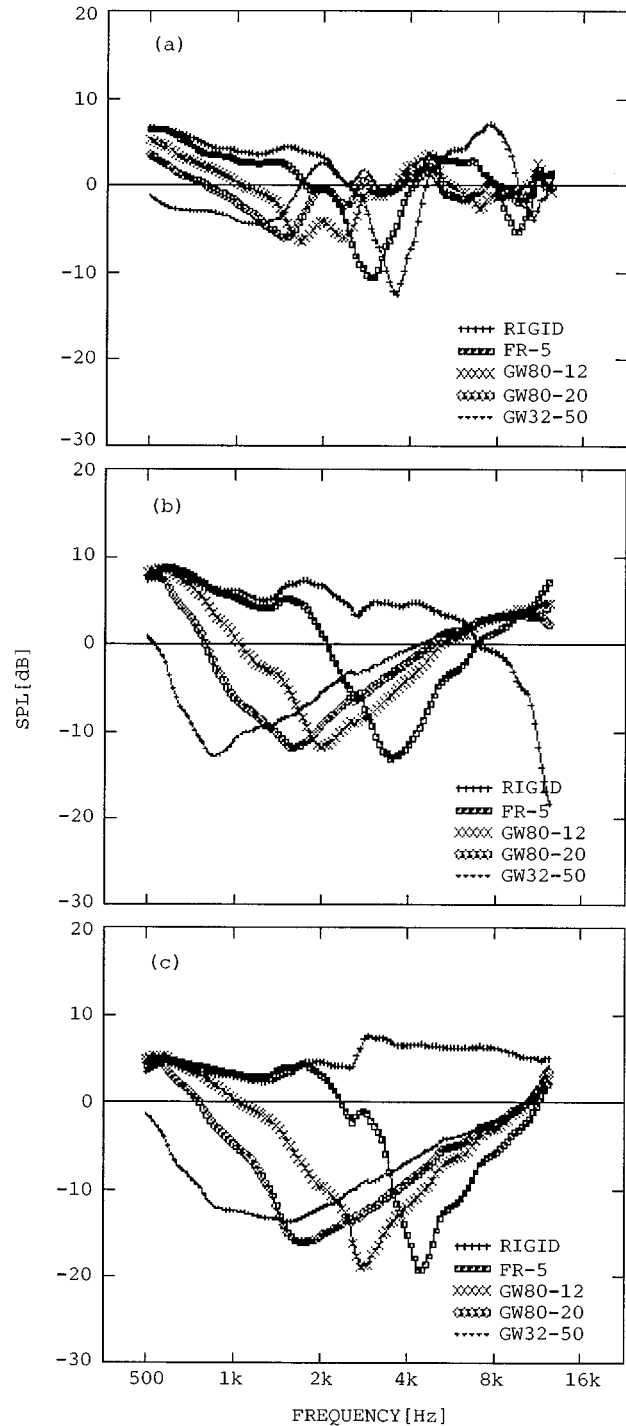
FIG. 12. Attenuation characteristics caused by various surface treatments at an angle of incidence (a) $\theta = 60°$, (b) $\theta = 80°$, and (c) $\theta = 85°$.

the absorptive treatment of seats and an audience as well as the floor is not always a fundamental solution to the problem of the SDE.

As long as an auditorium includes an audience, the seat dip attenuation is inevitable, and thorough improvement seems to be impossible. However, it may be expected that some measures for minimizing the effect will be found. The findings given by this investigation into the fundamental mechanism of the SDE will give an effective measure for this purpose.
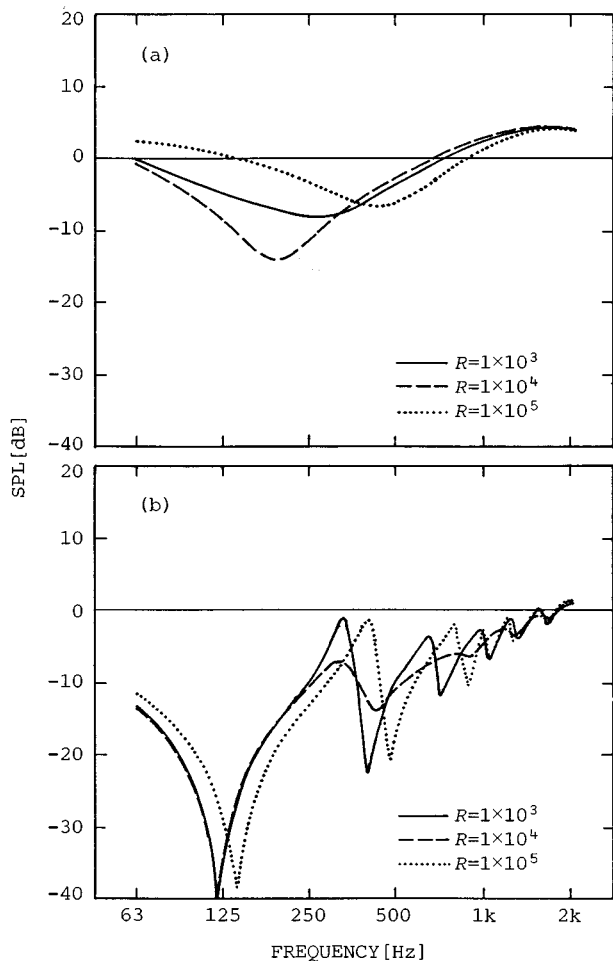
FIG. 13. Effect of seat backs evaluated by using two models of (a) extended and (b) local reaction, at $\theta = 85°$, with a parameter of flow resistivity R (mks-rayl/m).

FIG. 14. Effect of change in absorption of the seat surface, (a) SDE characteristics, and (b) seat absorption, at $\theta = 85°$. —, GW32-50; ---, GW32-100; ---; GW32-5 backed by an air layer of 500 mm.

IV. CONCLUSIONS

Experimental investigation using a $\frac{1}{10}$-scale model of a small hall was carried out. From the results of this investigation, it was confirmed that the phenomenon of seat dip attenuation always occurs even at rather late arriving sound. The degree of attenuation decreases as the arriving sound becomes later. In the present results, at least up to 16 ms (160 ms in full scale), a maximum attenuation of about -7 dB (relative to the case without seats) was observed.

To clarify the fundamental mechanism of the seat dip effect, a simple analytical model for a sound field caused by seat rows was derived from a wave scattering theory. In relation to this model, the reflected field from a layered boundary surface was analyzed by using a simple wave-reflection theory. This gives many clues about the fundamental mechanism of the SDE. From this analytical result and its experimental verification, it became clear that when the reflecting surface has any acoustic properties, except in the case of a perfectly rigid plane, the excess attenuation inevitably occurs, and the SDE is mainly characterized by seat backs or the upper part of human bodies.

The presented analytical model does not, however, cope with the phenomenon of primary dip-frequency shift with the

vertical angle of incidence, which may be caused by two-dimensional wave behavior in the seat rows. More rigorous theoretical models[8,16] can follow this phenomenon. However, it seems difficult, by using these models, to explain the cause for the phenomenon, because of a lack of any explicit analytical solutions for the models.

For controlling the effect, some possibilities were identified, as follows: The frequency of maximum dip and the degree of attenuation are exclusively determined by the admittance of a virtual surface at seat tops as well as the receiver height above the surface. Accordingly, some value of the admittance which can minimize the attenuation effect can presumably be found. The related parameters are the shape, structure, and material of the seat, seat arrangement on the floor, floor shape and its material, and so on. Further investigations will be required for these subjects, as well as the shift of primary dip with the vertical angle of incidence, the effect of horizontal angle of incidence, underpass transmission, and the effect on our auditory sense.

[1] T. J. Shultz and B. G. Watters, "Propagation of sound across audience seating," J. Acoust. Soc. Am. **36**, 885–896 (1964).
[2] G. M. Sessler and J. E. West, "Sound transmission over theatre seats," J. Acoust. Soc. Am. **36**, 1725–1732 (1964).

1333   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

D. Takahashi: Seat dip effect   1333

[3] K. Ishida, ''The measurement and prediction of sound transmission over auditorium seats,'' Ph.D. thesis, Cambridge University, 1993.

[4] M. Barron, ''Bass sound in concert auditoria,'' J. Acoust. Soc. Am. **97**, 1088–1098 (1995).

[5] J. S. Bradley, ''Some further investigations of the seat dip effect,'' J. Acoust. Soc. Am. **90**, 324–333 (1991).

[6] L. Cremer and H. A. Müller, *Principles and Applications of Room Acoustics* (Applied Science Publishers, London, 1982), Vol. I, Chap. I.6.

[7] Y. Ando, M. Takaishi, and K. Tada, ''Calculations of the sound transmission over theater seats and method for its improvement in the low-frequency range,'' J. Acoust. Soc. Am. **72**, 443–448 (1982).

[8] Y. Kawai and T. Terai, ''Calculation of sound fields over audience seats by using integral equation method,'' Trans. ASME, J. Vib. Acoust. **113**, 22–27 (1991).

[9] Lord Rayleigh, ''On the dynamical theory of grating,'' Proc. R. Soc. London, Ser. A **79**, 399–416 (1907).

[10] M. E. Delany and A. Bazley, ''Acoustical properties of fibrous absorbent materials,'' Appl. Acoust. **3**, 105–116 (1970).

[11] P. M. Morse, *Vibration and Sound* (McGraw–Hill, New York, 1948).

[12] K. Ishida, ''Investigation of the fundamental mechanism of the seat-dip effect—Using measurements on a parallel barrier scale-model,'' J. Acoust. Soc. Jpn. (E) **16**, 105–114 (1995).

[13] Y. Sakurai, H. Morimoto, and K. Ishida, ''The reflection of sound at grazing angles by auditorium seats,'' Appl. Acoust. **39**, 209–227 (1993).

[14] I. Rudnick, ''The propagation of an acoustic wave along a boundary,'' J. Acoust. Soc. Am. **19**, 348–356 (1947).

[15] 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).

[16] J. LoVetri, D. Mardare, and G. Soulodre, ''Modeling of the seat dip effect using the finite-difference time-domain method,'' J. Acoust. Soc. Am. **100**, 2204–2212 (1996).

1334   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

D. Takahashi: Seat dip effect   1334

Propagation of transverse waves on piezoelectric plates supporting single or double side metal strip gratings

S. Ballandras, J. B. Briot, E. Gavignet, and E. Bigler
Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S., Associé à l'Université de Franche-Comté-Besançon, Institut des Microtechniques de Franche-Comté, 32 avenue de l'Observatoire 25044 Besancon Cedex, France

A theoretical model of transverse waves propagating on piezoelectric plates of finite thickness under metal strip gratings is proposed. Different mechanical and electrical boundary conditions are considered and their influence on the propagation characteristics of the waves is emphasized. In all the cases, the dispersion curve relating the angular frequency to the wave number exhibits more than only one stopband as found in the usual analysis of surface transverse waves on semi-infinite substrates. This model has been used to calculate the resonance frequencies of synchronous devices on AT-cut and Z-cut quartz plates. These theoretical predictions are compared to experimental results obtained using synchronous devices built on 128-μm-thick Z-cut plate of quartz and AT-cut plate of quartz and operating at frequency range 80–150 MHz. © *1997 Acoustical Society of America.* [S0001-4966(97)05308-3]

INTRODUCTION

The excitation and propagation of surface transverse waves (STW) on semi-infinite piezoelectric substrates have been widely studied[1–4] to develop high frequency resonators[5,6] and filters.[7,8] STW can be excited using classical interdigital transducers and are trapped near the surface by using a periodic grating composed either of metal strips or grooves machined at the propagation surface. They can only propagate along particular crystal orientation like the rotated Z axis of singly rotated quartz substrates.

Nevertheless, experimental and theoretical work[9,10] has shown the usefulness of STW on AT-cut of quartz, presenting simultaneously a high velocity surface propagation (close to 5100 m/s) and a first order temperature compensation which are major characteristics for the design of high frequency resonators and filters.

Contrary to Rayleigh waves which are very strongly attenuated along the substrate thickness, STW penetrate deeply into the substrate, with an exponential amplitude decay of characteristic length equal to 35 acoustic wavelengths λ_a[11] ($2\lambda_a$ in the case of Rayleigh waves). However, the analysis of the influence of the substrate thickness on STW properties has not received significant interest. Auld *et al.* have proposed a theoretical approach for modelling the STW resonances on corrugated isotropic plates of finite thickness.[12] Ten years after, Danicki *et al.* have also proposed a similar analysis,[13] even extended to double side corrugated isotropic plates, but none of these studies has provided quantitative results for the design of STW devices on crystal plates of finite thickness.

The study presented in this paper is devoted to developing a theoretical model of transverse wave propagation on piezoelectric plate of finite thickness, trapped by periodic gratings on one or both sides of the plate. This work was motivated by the improvement of our previous STW model,[14] but also by the possibilities to fabricate transverse

wave devices on thin plates of quartz (AT-cut and Z-cut) in order to validate the proposed theoretical approach.

Various combinations of boundary conditions have been investigated to point out their influence on the physical properties of transverse waves. These theoretical points are detailed in the first part of the paper.

A transmission line approach close to the one already developed for STW[15] has also been implemented to predict the resonance frequencies of the studied devices. The corresponding calculations are described in the following section.

Finally, experiments have been performed on Z-cut plates and AT-cut plates of quartz. Measured resonance frequencies are then compared to theoretically predicted ones, emphasizing the good agreement between both results. As a conclusion, possible applications of these new transverse wave devices are discussed.

I. THEORETICAL MODEL OF PROPAGATION

A. Propagation equations

A purely transverse propagation is considered in the present model. As in the case of STW on semi-infinite piezoelectric substrates, the mechanical and electrical parts of the wave are identically polarized according to the axis definition of Fig. 1. The propagation direction a_3 corresponds to the rotated Z axis of the plate. The only nonzero component of the mechanical field u_1 and the electrical potential ϕ are assumed to depend also on a_2. Since the wave is supposed to propagate under an infinite periodic grating along a_3 inducing periodic boundary conditions,[1,2] a Bloch–Floquet expansion will be used to express both mechanical displacement and electrical potential as follows:
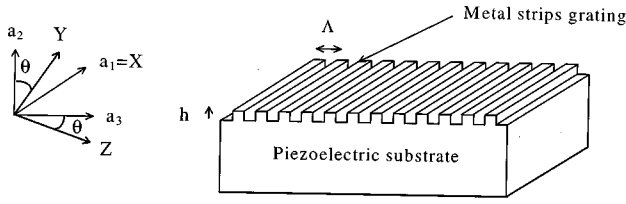
FIG. 1. The STW on piezoelectric plates of finite thickness: axis definition.

$$u_1 = \left(\sum_{n=-\infty}^{+\infty} A_n e^{-\alpha_n a_2} e^{-j(2\pi n a_3/\Lambda)} \right) e^{-j\beta_0 a_3} e^{j\omega t},$$

$$(1)$$

$$\phi = \left(\sum_{n=-\infty}^{+\infty} B_n e^{-\alpha_n a_2} e^{-j(2\pi n a_3/\Lambda)} \right) e^{-j\beta_0 a_3} e^{j\omega t},$$

where β_0 is the wavenumber of the zeroth mode and Λ the grating period. An equivalent wavenumber β_n is generally defined[1,4] for the nth mode as follows:

$$\beta_n = \beta_0 + 2\pi n/\Lambda, \quad n = 0, \pm 1, \pm 2, \dots . \quad (2)$$

In Eqs. (1), A_n and B_n are respectively the mechanical and electrical amplitude of the nth mode, α_n is representative of the wave distribution along the thickness of the substrate (a_2) and ω is the angular frequency. Since the grating is also assumed infinite along a_1, u_1 and ϕ do not depend on this variable.

As in the usual cases of wave propagation in piezoelectric materials, the present model requires the acoustic propagation equation together with Poisson's condition expressed for an insulating medium. Considering the wave polarization of Eqs. (1) yields the following expressions of the propagation equation and Poisson's condition :

$$\rho_0 \frac{d^2 u_1}{dt^2} = C_{55} \frac{\partial^2 u_1}{\partial a_3^2} + 2C_{56} \frac{\partial^2 u_1}{\partial a_2 \partial a_3} + C_{66} \frac{\partial^2 u_1}{\partial a_2^2}$$
$$+ e_{35} \frac{\partial^2 \phi}{\partial a_3^2} + (e_{36} + e_{25}) \frac{\partial^2 \phi}{\partial a_2 \partial a_3} + e_{26} \frac{\partial^2 \phi}{\partial a_2^2}, \quad (3)$$

$$e_{35} \frac{\partial^2 u_1}{\partial a_3^2} + (e_{36} + e_{25}) \frac{\partial^2 u_1}{\partial a_2 \partial a_3} + e_{26} \frac{\partial^2 u_1}{\partial a_2^2} - \varepsilon_{33} \frac{\partial^2 \phi}{\partial a_3^2}$$
$$- 2\varepsilon_{23} \frac{\partial^2 \phi}{\partial a_2 \partial a_3} - \varepsilon_{22} \frac{\partial^2 \phi}{\partial a_2^2} = 0, \quad (4)$$

where C_{IJ} and e_{iJ} represent the elastic and piezoelectric constants in contracted indexes, ε_{ij} the dielectric coefficients, and ρ_0 the mass density of the substrate. Considering β_0 and ω as varying parameters of the problem and inserting Eqs. (1) in Eqs. (3) and (4) provide a compatibility condition between these equations. This condition is written as a fourth order polynomial in α_n:
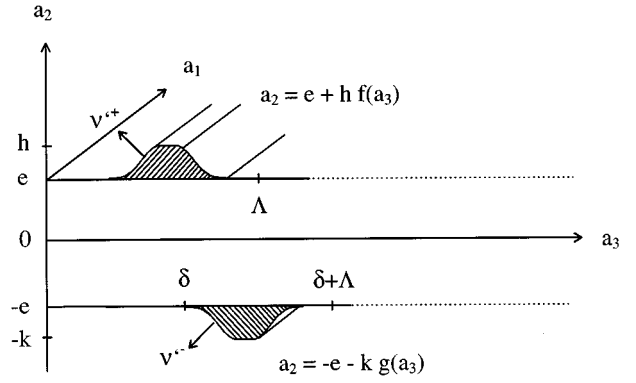
FIG. 2. Double side grating device built on piezoelectric plate, definition of the normal vectors pointing outward the metal strips.

$$\alpha_n^4 (C_{66}\varepsilon_{22} + e_{26}^2) + \alpha_n^3 2j\beta n (C_{66}\varepsilon_{23} + C_{56}\varepsilon_{22} + e_{26}(e_{36} + e_{25}))$$
$$- \alpha_n^2 \beta_n^2 (\varepsilon_{22}C_{55} + \varepsilon_{33}C_{66} + 4\varepsilon_{23}C_{56} - \varepsilon_{22}(\rho_0\omega^2/\beta_n^2)$$
$$+ (e_{36} + e_{25})^2 + 2e_{26}e_{35})$$
$$- \alpha_n 2j\beta_n^3 (C_{56}\varepsilon_{33} + C_{55}\varepsilon_{23} - \varepsilon_{23}(\rho_0\omega^2/\beta_n^2)$$
$$+ e_{35}(e_{36} + e_{25}))$$
$$+ \beta_n^4 (C_{55}\varepsilon_{33} + e_{35}^2 - \varepsilon_{33}(\rho_0\omega^2/\beta_n^2)) = 0. \quad (5)$$

As a consequence, four values of α_n have to be considered for each mode n. In the case of semi-infinite substrate, only two of these values correspond to a physically realistic solution of a wave decreasing along a_2 (real part of α_n negative). In the present case, physical considerations cannot be invoked to suppress any value of α_n, and wave components of Eqs. (1) are then composed of four partial waves written as follows:

$$u_1 = \sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} A_{nr} e^{-\alpha_{nr}a_2} \right) e^{-j\beta_n a_3} e^{j\omega t},$$

$$(6)$$

$$\phi = \sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} B_{nr} e^{-\alpha_{nr}a_2} \right) e^{-j\beta_n a_3} e^{j\omega t}.$$

Considering now Poisson's condition (4) together with Eqs. (6) allows us to relate mechanical and electrical amplitudes as follows:

$$B_{nr} = D_{nr} A_{nr}$$

$$\text{with} \quad D_{nr} = \frac{-e_{35}\beta_n^2 + j(e_{36} + e_{25})\alpha_{nr}\beta_n + e_{26}\alpha_{nr}^2}{-\varepsilon_{33}\beta_n^2 + 2j\varepsilon_{23}\alpha_{nr}\beta_n + \varepsilon_{22}\alpha_{nr}^2}. \quad (7)$$

This relation allows us to reduce the number of independent unknowns for each mode. These unknowns are now determined using boundary conditions.

B. Mechanical boundary conditions

We consider first the case of a double side grating device as shown in Fig. 2. The case of single side grating device (Fig. 3) represents only a simplification of this more complete one, as it is shown further. For more generality, the top and bottom strip profiles are assumed different, with respective maximum heights h and k. In the present analysis,
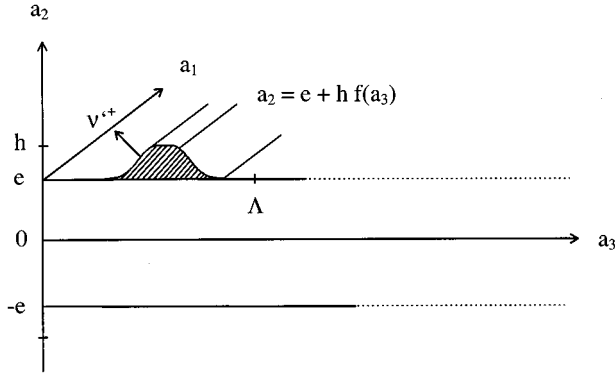
FIG. 3. Single side grating device. The bottom surface may be considered electrically shorted or open.

only gratings of identical spatial period are considered in order to respect the validity of the Bloch–Floquet expansion of Eq. (1).

However, the model takes into account a possible shift δ between the origins of both gratings (cf. Fig. 2). The mechanical boundary conditions are then established in two steps. The first step consists in expressing a stress free condition at the outer surface of the strips. Assuming that the strip profile can be expressed as a function of the propagation axis a_3,[15] this stress free condition is written as follows for the top and bottom side strips:

$$T'_{ij}\nu'^+_j = 0 \quad \text{along} \quad a_2 = e + h \cdot f(a_3),$$

$$T'_{ij}\nu'^-_j = 0 \quad \text{along} \quad a_2 = -e - k \cdot g(a_3). \tag{8}$$

In Eqs. (8), e represents the half-thickness of the piezoelectric plate, and $f(a_3)$ and $g(a_3)$ are respectively the profile functions of top and bottom side strips. The propagation characteristics and mechanical properties relative to the strip are primed. T'_{ij} is the dynamical stress tensor, ν'^+_j and ν'^-_j are the normal vectors pointing respectively outwards of the top and bottom side strip surfaces as shown in Fig. 2. According to Ref. 14 the component of ν'^+_j and ν'^-_j are written as follows:

$$\nu'^+_1 = 0, \quad \nu'^-_1 = 0, \quad \nu'^+_2 = 1, \quad \nu'^-_2 = -1,$$

$$\nu'^+_3 = -h\frac{\partial f(a_3)}{\partial(a_3)}, \quad \nu'^-_3 = -k\frac{\partial g(a_3)}{\partial(a_3)}. \tag{9}$$

The second step of the development of the mechanical boundary conditions consists in establishing continuity relations at the interface between the metal strip and the piezoelectric medium. These relations which concern the mechanical displacements and the nonzero normal stress components are written on the top and bottom surfaces of the piezoelectric plate:

$$u'_1(-e,a_3) = u_1(-e,a_3), \quad T'_{12}(-e,a_3) = T_{12}(-e,a_3), \tag{10}$$

$$u'_1(e,a_3) = u_1(e,a_3), \quad T'_{12}(e,a_3) = T_{12}(e,a_3).$$

In Eqs. (10), the mechanical displacements in the plate are obtained from Eq. (6) and the stress component T_{12} is given by the piezoelectric Hooke's law combined with the mechanical displacement and the electrical potential of Eq. (6):

$$T_{12}(a_2,a_3) = -\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}[(C_{66}\alpha_{nr}+jC_{56}\beta_n)+(e_{26}\alpha_{nr}$$
$$+je_{36}\beta_n)A_{nr}e^{-\alpha_{nr}a_2}]e^{-j\beta_n a_3}e^{j\omega t}. \tag{11}$$

The expression of T'_{12} is now derived from Eqs. (8) and (9). Considering that the strip height is much smaller than the grating period Λ allows a Taylor–Maclaurin expansion of the stress condition of Eq. (8). When limited to the first order along the strip height h or k, this expansion directly provides a simple expression of $T'_{12}(e,a_3)$ and $T'_{12}(-e,a_3)$:

$$T'_{12}(e,a_3) = -h\left(f(a_3)\frac{\partial T'_{12}(e,a_3)}{\partial a_2}\right.$$
$$\left. -\frac{\partial f(a_3)}{\partial a_3}T'_{13}(e,a_3)\right).$$
$$T'_{12}(-e,a_3) = k\left(g(a_3)\frac{\partial T'_{12}(-e,a_3)}{\partial a_2}\right.$$
$$\left. -\frac{\partial g(a_3)}{\partial a_3}T'_{13}(-e,a_3)\right). \tag{12}$$

A correct expression of the derivatives along a_2 of Eqs. (12) requires an analytical formula of u'_1. Considering the continuity relations (10), this analytical formula is written as follows:

$$u'_1 = \sum_{n=-\infty}^{+\infty} G'_n(a_2)e^{-j\beta_n a_3}e^{j\omega t},$$

$$\text{with} \quad G'_n(e) = \sum_{r=1}^{2} A_{nr}e^{-\alpha_{nr}e},$$

$$\text{and} \quad G'_n(-e) = \sum_{r=1}^{2} A_{nr}e^{+\alpha_{nr}e}. \tag{13}$$

The function $G'_n(a_2)$ of Eq. (13) provides the behavior of u'_1 along a_2 and is determined using the propagation equation for an isotropic medium. However, the complete determination of $G'_n(a_2)$ is not necessary since this propagation equation provides a substitution expression of $\partial T'_{12}/\partial a_2$ as proposed below:

$$\frac{\partial T'_{12}}{\partial a_2} = -\left(\rho'\omega^2 u'_1 + \frac{\partial T'_{13}}{\partial a_3}\right) \quad \text{with} \quad T'_{13} = C'_{55}\frac{\partial T'_{13}}{\partial a_3}, \tag{14}$$

where ρ' is the mass density of the metal of the strip and C'_{55} its shear elastic constant. According to Eqs. (10), (11), (12), (13), and (14), the mechanical boundary conditions are given by the following expressions:

$$\sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} [C_{66}\alpha_{nr} + jC'_{56}\beta_n + D_{nr}(e_{26}\alpha_{nr} \right.$$

$$+ je_{36}\beta_n)]A_{nr}e^{-\alpha_{nr}e}e^{-j\beta_n a_3}$$

$$= -h\sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} \left[(\rho'\omega^2 - C'_{55}\beta_n^2)f(a_3) \right.\right.$$

$$\left.\left. - jC'_{55}\beta_n\frac{\partial f(a_3)}{\partial a_3} \right]A_{nr}e^{-\alpha_{nr}e} \right)e^{-j\beta_n a_3},$$

$$(15)$$

$$\sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} [C_{66}\alpha_{nr} + jC_{56}\beta_n + D_{nr}(e_{26}\alpha_{nr} \right.$$

$$+ je_{26}\beta_n)]A_{nr}e^{+\alpha_{nr}e}e^{-j\beta_n a_3}$$

$$= +k\sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} \left[(\rho'\omega^2 - C'_{55}\beta_n^2)g(a_3) \right.\right.$$

$$\left.\left. - jC'_{55}\beta_n\frac{\partial g(a_3)}{\partial a_3} \right]A_{nr}e^{+\alpha_{nr}e} \right)e^{-j\beta_n a_3}.$$

Solving Eqs. (15) requires the elimination of their dependence along a_3. This can be performed by integrating these equations over one period Λ along a_3, since the gratings are assumed infinite along this direction. However, it is relevant to take advantage of the orthogonal properties of the different modes as proposed by Auld.[1]

Thus, each part of Eq. (15) is multiplied by $e^{-j\beta_q a_3}$ and the integration along a_3 over one period Λ is then performed. This procedure is equivalent to simplify Eqs. (15) by suppressing on each side the term $e^{-j\beta_0 a_3}$ and performing an orthogonal projection of the resulting equations over the function base $e^{+j(2\pi q/\Lambda)}$. It provides the following expressions of the mechanical boundary conditions:

$$\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[K_{qnr} + \frac{\Lambda}{h}\delta_{nq}(F_{nr} + D_{nr}Z_{nr}) \right]e^{-\alpha_{nr}e}A_{nr} = 0,$$

$$q = 0, \pm 1, \pm 2,\dots ,\quad (16)$$

$$\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[L_{qnr} - \frac{\Lambda}{k}\delta_{nq}(F_{nr} + D_{nr}Z_{nr}) \right]e^{+\alpha_{nr}e}A_{nr} = 0,$$

with

$$K_{qnr} = (\rho'\omega^2 - C'_{55}\beta_n\beta_q)I_{nq},$$

$$L_{qnr} = (\rho'\omega^2 - C'_{55}\beta_n\beta_q)J_{nq},$$

$$F_{nr} = C_{66}\alpha_{nr} + jC_{56}\beta_n, \quad Z_{nr} = e_{26}\alpha_{nr} + je_{36}\beta_n,$$

$$I_{nq} = \int_{-\Lambda/2}^{\Lambda/2} f(a_3)e^{-j(\beta_n - \beta_q)a_3}da_3,$$

$$J_{nq} = \int_{\delta-\Lambda/2}^{\Lambda/2+\delta} g(a_3)e^{-j(\beta_n - \beta_q)a_3}da_3,$$

where δ represents the shift between the two grating origins, conforming to Fig. 2. In the case of a single side grating structure, the boundary conditions (16) can be directly used by setting L_{qnr} equal to zero.

C. Electrical boundary conditions

The rigorous development of electrical boundary conditions of a surface wave device based on interdigital transducers and metal strip gratings may require complicated models, particularly for the calculation of the electrical charge distribution at the interface between the metal and the substrate. However, for weakly coupled piezoelectric substrates like quartz, simple approximations can be applied without strong losses of accuracy. More, it was shown in the case of STW on quartz[15] that such an approach provides significant improvement in the prediction of resonance frequencies compared to a nonpiezoelectric analysis and experimental results.

These approximations correspond to the two following academic problems. The first consists in considering a propagation surface covered with a metal layer of negligible mass and electrically shorted. The metal strip grating is assumed to be deposited on this (ideally thin but perfectly conductive) metal layer. The second simplified case taken into consideration is based on a propagation surface free of metallization but supporting hypothetic dielectric strip grating of the same mechanical properties as practically used for real metal strips. Both approaches only take into account the mass loading effect due to the grating. However, they allow us to introduce piezoelectricity in the propagation equations and boundary conditions in order to provide better results than a simple nonpiezoelectric mechanical model in regard with experiments. This approach also provides simpler expressions (very easy to compute) than a complete electrical model does.

In the case of a grounded propagation surface, the electrical boundary conditions are written:

$$\sum_{r=1}^{4} D_{qr}A_{qr}e^{-\alpha_{qr}e} = 0 \quad \text{along } a_2 = +e,$$

$$q = 0, \pm 1, \pm 2,\dots , \quad (17)$$

$$\sum_{r=1}^{4} D_{qr}A_{qr}e^{+\alpha_{qr}e} = 0 \quad \text{along } a_2 = -e.$$

In the case of a dielectric strip grating, the continuity of the electrical potential and the normal component of the electrical displacement vector must be established:

$$\left.\begin{array}{c} D_2^s - D_2^v = 0 \\ \phi^s - \phi^v = 0 \end{array}\right\} \quad \text{along } a_2 = \pm e, \quad (18)$$

$$\phi^v = 0 \quad \text{along } a_2 = \pm\infty,$$

where the superscripts s and v denote respectively the substrate and the vacuum. Conditions (18) require the calculation of the potential distribution in vacuum. This distribution is derived from Poisson's condition in vacuum:

$$\frac{\partial^2 \phi^v}{\partial a_2^2} + \frac{\partial^2 \phi^v}{\partial a_3^2} = 0. \quad (19)$$

The dependence of ϕ^v along a_3 is given by the continuity relations (18) which also provides the amplitude of the potential in vacuum. As for mechanical displacements in the strip, the dependence of ϕ^v along a_2 must be determined. In

this matter, the following form of ϕ^v is considered:

$$\phi^v = \sum_{n=-\infty}^{+\infty} G_n^v(a_2)e^{-j\beta_n a_3}e^{j\omega t},$$

$$\text{with } G_n^v(\pm e) = \sum_{r=1}^{4} D_{nr}A_{nr}e^{\mp \alpha_{nr}e}. \qquad (20)$$

From (19) and (20), the expression of $G_n^v(a_2)$ is derived as follows:

$$G_n^v(a_2) = A_n e^{\beta_n a_2} + B_n e^{-\beta_n a_2}. \qquad (21)$$

The constants A_n and B_n are then determined using both Eq. (20) and the condition of potential vanishing at $+$ or $-$ infinity. Finally, the expression of the potential in vacuum is given as follows:

$$\phi^v = \sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} D_{nr}A_{nr}e^{-\alpha_{nr}e} \right) e^{-|\beta_n|(a_2-e)}e^{-j\beta_n a_3}e^{j\omega t}$$

along a_2 greater than or equal to e,

$$\phi^v = \sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} D_{nr}A_{nr}e^{+\alpha_{nr}e} \right) e^{|\beta_n|(a_2+e)}e^{-j\beta_n a_3}e^{j\omega t}$$

$$(22)$$

along a_2 lower than or equal to $-e$.

As a consequence, the continuity of the electrical displacement normal to the plate is now written under its final form:

$$\sum_{r=1}^{4} [\alpha_{qr}e_{26} + j\beta_q e_{36}D_{qr}(\varepsilon_{22}\alpha_{qr} + j\varepsilon_{23}\beta_q - |\beta_q|\varepsilon_0)]$$

$$\times A_{qr}e^{-\alpha_{qr}e} = 0 \quad \text{along } a_2 = +e,$$

$$\sum_{r=1}^{4} [\alpha_{qr}e_{26} + j\beta_q e_{36} - D_{qr}(\varepsilon_{22}\alpha_{qr} + j\varepsilon_{23}\beta_q + |\beta_q|\varepsilon_0)]$$

$$(23)$$

$$\times A_{qr}e^{+\alpha_{qr}e} = 0 \quad \text{along } a_2 = -e.$$

The electrical boundary conditions may also result of a mixing of conditions (17) and (23). For example, a single side grating device can be considered as grounded on its top surface, and free of metallization on its bottom side. It is then relevant to use Eq. (17) along $a_2 = +e$ and Eqs. (23) along $a_2 = -e$.

D. Numerical results

The proposed model has been written in order to respect the Bloch–Floquet theorem which requires infinite sums of modes for mechanical displacements and electrical potentials (since the gratings have been assumed infinite along a_3). However, it is quite obvious that this rigorous approach is not adapted to effective computing operations. As a consequence, truncating these sums is required to implement the model using computer programs. In the case of STW propagating on semi-infinite substrate under a metal strip of height much smaller than the grating period Λ, a coupled mode approach[1,2,5,6,14] is sufficient to correctly predict the properties of the waves (when compared to experimental measurements). This approach has then been applied in the present analysis, retaining only modes $n=-1$ and $n=0$ in the pre-

viously developed equations. However, the possibility to extend the number of modes taken into account for computing has been integrated in the computer programs.

The last point to consider is the metal strip profile functions $f(a_3)$ and $g(a_3)$ which must be defined to perform integrals I_{nq} and J_{nq} of Eq. (16). Generally, technologies used to machine the metal strip gratings of acoustic wave devices do not allow a precise control of the shape of the strips. Moreover, this shape is very difficult to measure and may vary from one device to another. Nevertheless, many works have shown[1,2,5,6,15] that a rectangular profile function can be efficiently used to provide accurate predictions of the wave properties and characteristics. This kind of function has been chosen to solve the integrals of Eq. (16). It is based on a combination of heaviside functions H as follows:

$$f(a_3) = H\left(a_3 + \frac{r_t\Lambda}{2}\right) - H\left(a_3 \frac{r_t\Lambda}{2}\right),$$

$$g(a_3) = H\left(a_3 + \frac{r_b\Lambda}{2} - \delta\right) - H\left(a_3 - \frac{r_b\Lambda}{2} - \delta\right)$$

$$(24)$$

with $0 < r_t, r_b < 1$.

In Eqs. (24), r_t and r_b represent, respectively, the metallization ratio on the top and bottom sides of the plate. In the very classical case of a metallization ratio $r_{t,b} = \frac{1}{2}$, the integrals I_{nq} and J_{nq} take the following values:

$$I_{nq} = \frac{\Lambda}{2}\sin\left(\frac{\pi}{2}(n-q)\right) \Big/ \frac{\pi}{2}(n-q),$$

and

$$(25)$$

$$J_{nq} = e^{j(\beta_n - \beta_q)\delta}I_{nq}, \quad 0 < \delta < \Lambda.$$

For practical computations, it was necessary to fix all the variable parameters of the model. The simplest configuration consisting in setting $h=k$, $f(a_3) = g(a_3)$, $r_t = r_b = \frac{1}{2}$ and $\delta = 0$ have been chosen first. Singly rotated quartz plates have been considered as propagation substrate. Cut angles ϕ and θ are defined according IEEE 49 Standard.[16] The elastic, piezoelectric, and dielectric coefficients of quartz were given by Slobodnik.[17] The period of the gratings Λ and strip height h have been respectively set to 20 μm and 1500 Å and the thickness of the quartz plate was varied from 130 to 150 μm to check its influence on the wave properties. The metal of the strip was aluminum with a mass density set to 2695 kg m^{-3} and a shear elastic constant C'_{55} equal to 25 GPa. Calculations were performed for both AT-cut ($\theta = 36°$) and Z-cut (θ close to 90°) of quartz corresponding to plates available on the market. An algorithm very close to the one used for classical STW was used.[14]

Three kinds of boundary conditions have been tested to compare respective properties of each configuration. These properties are mainly illustrated by the dispersion relation which relates β_0 to ω. Figure 4 shows this dispersion relation in the case of an AT-cut plate of quartz supporting a double side grating structure, for different values of thickness $2e$.

The most significant result emphasized by the curves of Fig. 4 is the multiplicity of frequency stopbands, correspond-
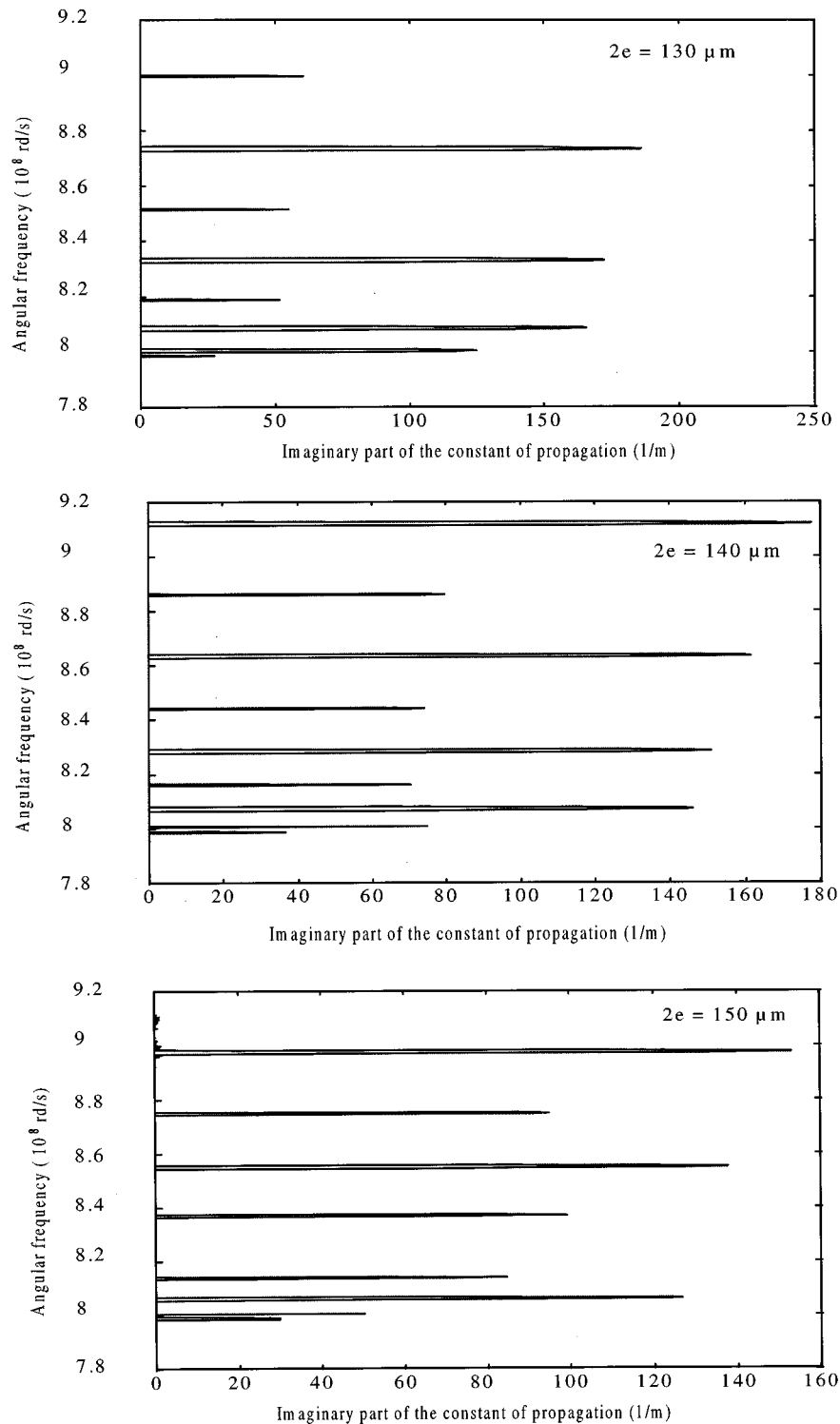
FIG. 4. Dispersion relations (imaginary part) between ω and β_0 for double side grating device on AT cut of quartz for different values of the plate thickness $2e$. (a) $2e = 130$ μm, (b) $2 e = 140$ μm, (c) $2 e = 150$ μm.

ing to a nonzero imaginary part of the propagation coefficient β_0. The first stopband is the same as the one found for STW on semi-infinite substrate (real part of β_0 equal to π/Λ). However, after this first stopband, β_0 becomes again purely real. This signifies that waves can propagate at higher frequencies without being reflected by the strips of the gratings. Since β_0 is purely real outside the stopbands, multiple

resonances are expected for a given design. This point is developed in the following section of this paper.

The comparison between the three curves of Fig. 4 indicates that the number of frequency stopbands is strongly related to the thickness of the substrate. The number of predicted stopbands increases together with the thickness of the plate. A slight shift of the frequency stopbands can also be
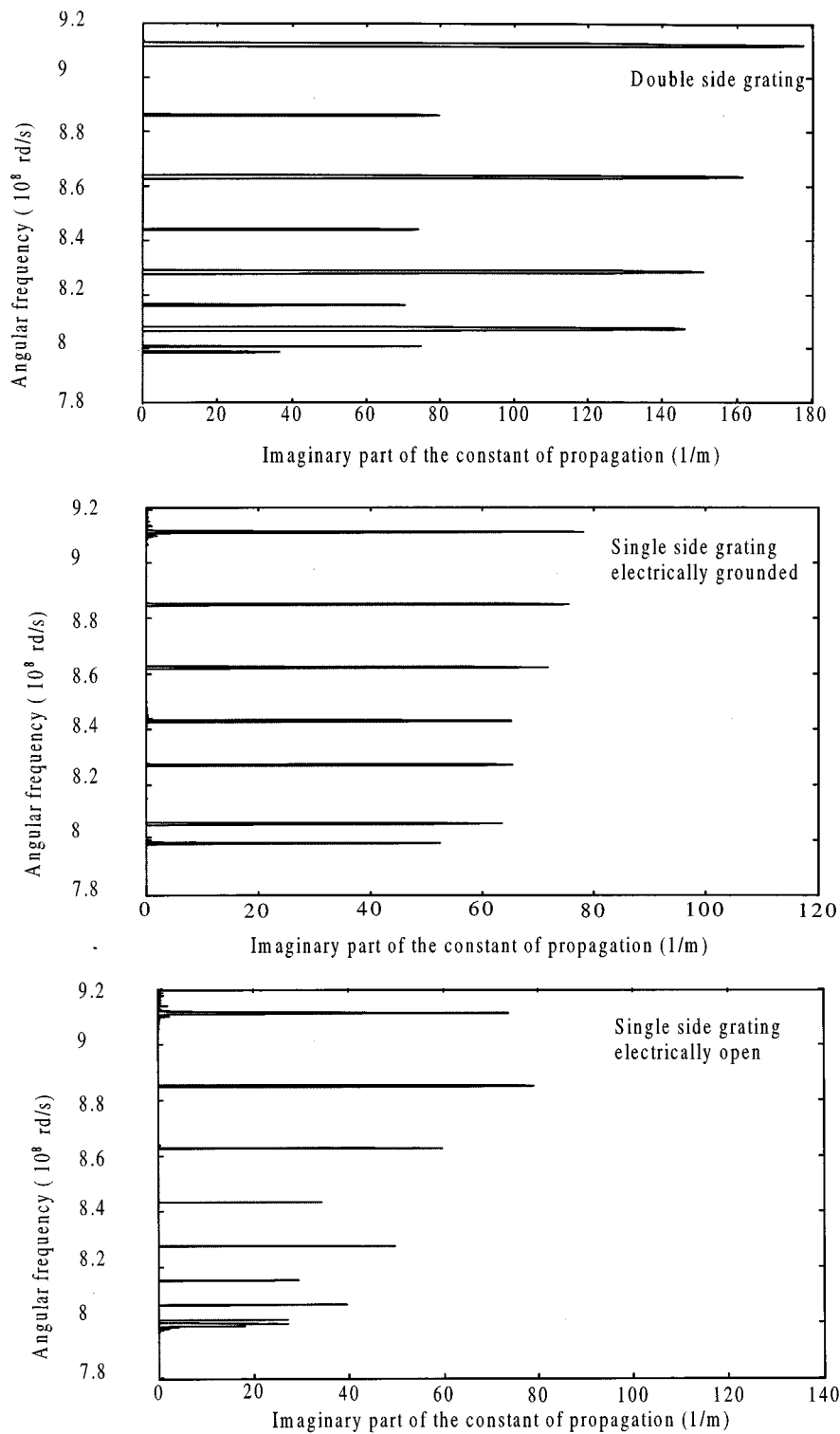
FIG. 5. Dispersion relations (imaginary part) compared for different boundary conditions, case of on AT cut quartz plate of thickness $2e = 140$ μm. (a) Double side grating device, (b) single side grating device, bottom side electrically grounded, and (c) single side grating device, bottom side electrically open.

noted when comparing the three dispersion curves. An explanation can be proposed by comparing this phenomenon to the excitation of plate modes propagating on singly rotated quartz plates.[18-20] Such a comparison also allows us to explain the multiple frequency stopbands of Fig. 4. In the classical plate mode propagation, the component along a_3 is progressive but the one along a_2 is trapped between both sides of the plates. In the present case, both parts of the wave

(along a_2 and a_3) can be trapped due to the presence of the gratings, even for high order plate modes, as shown in Fig. 4.

Figure 5 shows a comparison of dispersion curves calculated for different boundary conditions on an AT-cut plate of quartz of thickness 140 μm. The curve (a) corresponds to the previous case of a double side grating structure, both of the sides being electrically grounded, while curves (b) and (c) correspond to a single side grating device with the non-

mechanically loaded side respectively electrically grounded [Eq. (17)] and free [Eq. (23)].

This comparison emphasizes two important properties of the corresponding devices. First, it can be noted that the value of the imaginary part of the propagation coefficient β_0 calculated for the double side grating structure is twice as high as the one for the single side design for some of the modes. This indicates that the corresponding waves are more rapidly attenuated along a_3 in double side grating structures than in single sided ones when operating in frequency stop-bands. The second point concerns the control of the number of stopbands using appropriate boundary conditions. Comparison of Fig. 5(b) and (c) shows that grounding the non-mechanically loaded side induces the vanishing of some stopbands. The same results and conclusions can be found considering a Z-cut plate of quartz. However, they have to be confirmed experimentally. In this approach, the proposed model must be completed to provide information easily comparable to those obtained using a network analyzer.

II. RESONANCE FREQUENCY CALCULATION

In this section, a transmission line model is proposed to calculate the resonance frequency of synchronous grating resonators.[6,8,15] This approach has been successfully developed to predict the resonance frequency of three cavity STW resonators[21] and is easily transposable to the present problem. A transfer matrix has to be defined first for any periodic grating. The combination of transfer matrices then provides resonance conditions which have to be satisfied by the wave propagating in the structure. These two steps are detailed in the following paragraphs.

A. Transfer matrix

The approach proposed here consists in determining a set of coefficients relating incident and reflected waves on each side of a periodic grating, as shown in Fig. 6. Incident waves are noted (u_{11}^+, u_{12}^-) and reflected waves are noted (u_{11}^-, u_{12}^+), according to the wave orientations defined in Fig. 6. The origin of the grating is assumed at any distance d of the one used to calculate the wave properties (Sec. I) and its length is L. Each wave is composed of direct and reverse progressive waves corresponding respectively to the first and second Floquet's mode, both solutions of the problem detailed in Sec. I. To simplify the equations, each mechanical displacement u_1 is expressed along $a_2=0$ (vanishing of the dependence along a_2). This assumption limits the relevance of the proposed approach to single side grating and symmetrical ($\delta=0$) double side grating devices. For more complicated structures, the dependence of the wave along a_2 has to be taken into account.
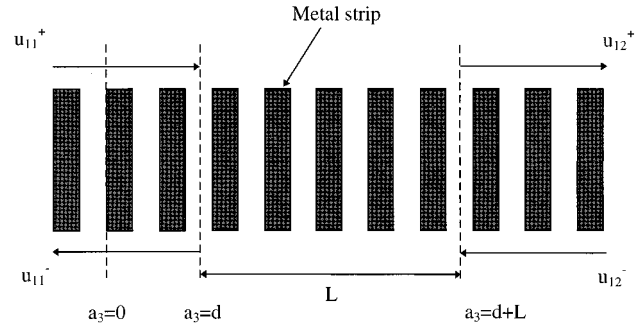
FIG. 6. Representation of a metal strip grating by a transfer matrix, definition of the wave orientations.

It is now necessary to express the second Floquet's mode to provide a complete description of the incident and reflected waves of Fig. 6. This is obtained by changing the propagation coefficient β_0 to $\pi/\Lambda - \beta_0$.[21] It is then possible to establish relations between the propagation characteristics for each mode, as shown in the Appendix. For more clarity, the first and second mode parameters are distinguished using superscripts st and nd. Considering now that β_0 can be replaced by $\pi/\Lambda + \delta\beta$ (where Re ($\delta\beta$) is always smaller than π/Λ) allows us to separate the mechanical displacement u_1 at port 1 of the grating of Fig. 6 in direct and reverse wave components as follows:

$$
\begin{aligned}
u_{11}^+ &= \sum_{n=0}^{+\infty} \left\{ \sum_{r=1}^{4} A_{nr}^{st} e^{-j((2n+1)\pi/\Lambda + \delta\beta)d} \right. \\
&\left. + \sum_{r=1}^{4} A_{nr}^{nd} e^{-j((2n+1)(\pi/\Lambda) - \delta\beta)d} \right\}, \\
u_{11}^- &= \sum_{n=-\infty}^{-1} \left\{ \sum_{r=1}^{4} A_{nr}^{st} e^{-j((2n+1)(\pi/\Lambda) + \delta\beta)d} \right. \\
&\left. + \sum_{r=1}^{4} A_{nr}^{nd} e^{-j((2n+1)(\pi/\Lambda) - \delta\beta)d} \right\}.
\end{aligned}
\tag{26}
$$

The expressions of u_{12}^+ and u_{12}^- are given by replacing d by $d+L$ in Eq. (26). A first matrix relation can be deduced from Eq. (26) by setting $A_{nr}^{st} = x_{nr}^{st} A_0^{st}$ and $A_{nr}^{nd} = x_{nr}^{nd} A_0^{nd}$:

$$
\begin{Bmatrix} u_{11}^+ \\ u_{11}^- \end{Bmatrix} = [P(d)] \begin{Bmatrix} A_0^{st} \\ A_0^{nd} \end{Bmatrix},
$$

$$
\begin{Bmatrix} u_{12}^+ \\ u_{12}^- \end{Bmatrix} = [P(d+L)] \begin{Bmatrix} A_0^{st} \\ A_0^{nd} \end{Bmatrix},
$$

where

$$
[P(a_3)] = \begin{bmatrix} \sum_{n=0}^{+\infty}\sum_{r=1}^{4} x_{nr}^{st} e^{-j((2n+1)(\pi/\Lambda)+\delta\beta)a_3} & \sum_{n=0}^{+\infty}\sum_{r=1}^{4} x_{nr}^{nd} e^{-j((2n+1)(\pi/\Lambda)-\delta\beta)a_3} \\ \sum_{n=-\infty}^{-1}\sum_{r=1}^{4} x_{nr}^{st} e^{-j((2n+1)(\pi/\Lambda)+\delta\beta)a_3} & \sum_{n=\infty}^{-1}\sum_{r=1}^{4} x_{nr}^{nd} e^{-j((2n+1)(\pi/\Lambda)-\delta\beta)a_3} \end{bmatrix}.
\tag{27}
$$
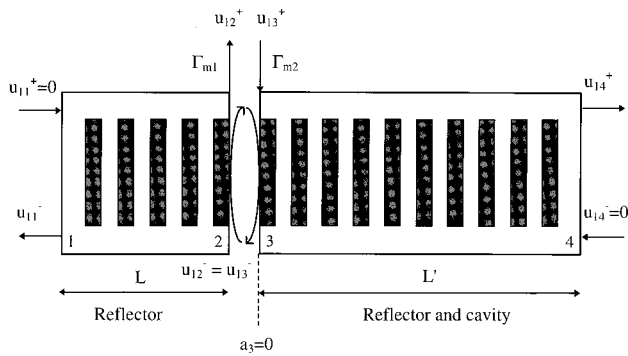
FIG. 7. Synchronous device composed of two grating structures of respective length L, L' (distance from the origin respectively equal to d and d').

The expression of the transfer matrix M is then directly provided by the following combination of matrix $P(d)$ and $P(d+L)$:

$$M = [P(d)][P(d+L)]^{-1}. \tag{28}$$

This transfer matrix can be used to combine different gratings and to develop resonance conditions, as explained in Ref. 15 in the case of asynchronous three-cavity devices.

B. Resonance condition

In the present study, we focused our interest on the case of synchronous devices. These devices can be modelled by combining two gratings of the same period Λ, as shown in Fig. 7. The interdigital transducers are taken into account in each grating, but only their acoustical contribution is considered. The only waves assumed to propagate in the structure are those excited by these transducers. Since no waves are entering the structure from the outside, the problem will be simplified by setting u_{11}^+ and u_{14}^+ to zero. This provides a relation between u_{12}^+ and u_{12}^-, and also between u_{13}^+ and u_{13}^-, written as follows:

$$u_{12}^+ = \Gamma_{M2} u_{12}^-, \quad u_{13}^+ = \frac{u_{13}^-}{\Gamma_{M3}}, \tag{29}$$

where Γ_{M2} and Γ_{M3} are the reflection coefficients at ports 2 and 3 of the structure of Fig. 7. Since port 2 and port 3 are directly connected together, u_{12}^+ and u_{13}^+ are equal, as are u_{12}^- and u_{13}^-. The resonance condition is then easily deduced from Eq. (29):

$$\Gamma_{M2} \cdot \Gamma_{M3} = 1. \tag{30}$$

The resonance condition of Eq. (30) is used to determine the angular frequency ω which appeared as a varying parameter in Sec. I. This calculation can be applied for all the modes propagating in the plate, corresponding to the different frequency stopbands previously emphasized. Theoretical results are compared to experimental data in the following section.

III. EXPERIMENTAL RESULTS

Experimental tests have been performed on a Z-cut plate and an AT-cut plate of quartz, of thickness 128 μm. Synchronous delay lines and resonators have been patterned on
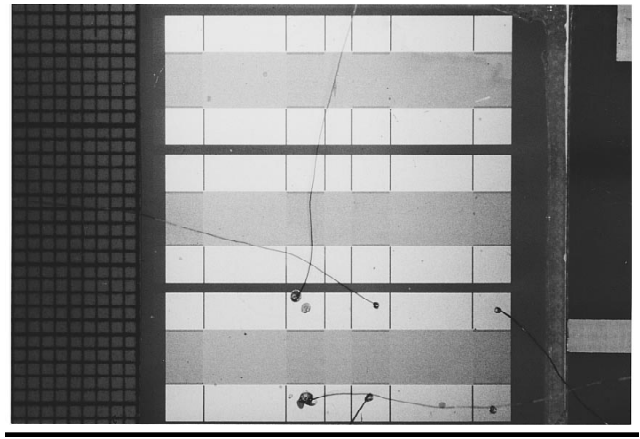
FIG. 8. Photograph of an experimental device placed in a metal box for electrical measurements.

one side of these plates, the other was left unloaded and ungrounded. One of the tested devices is shown in Fig. 8.

The width of the aluminum strips was 10 μm and their height was 150 nm. Each interdigital transducer was composed of fifty pairs of strips.

The transmission response of the central resonator structure has been measured using a network analyzer in the frequency range 80–150 MHz for Z-cut devices, and 120–170 MHz for AT-cut devices. Resonance frequencies are measured at a maximum of amplitude corresponding to a zero of phase. An example of electrical response is plotted in Fig. 9 in the case of an AT-cut plate. Insertion losses better than -10 dB and Q factor in excess of 25 000 have been measured at frequency close to 130 MHz. Comparison between measured frequencies and a theoretically predicted one is presented in Table I, emphasizing the very good agreement of both results for the AT-cut plate device. However, slight discrepancies can be pointed out between theory and experiments in the case of the Z-cut device. Since transverse waves propagating along the surface of this device present a polarization very close to thickness shear bulk waves in a Y-cut of quartz, they should also exhibit a larger piezoelectric cou-
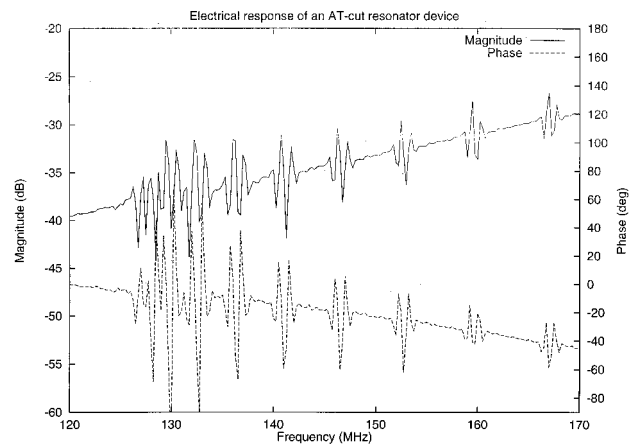
FIG. 9. Electrical response of an AT-cut place resonator-like device (plate thickness 128μm, strip width and height respectively equal to 10 μm and 150 nm).

Ballandras *et al.*: Transverse waves on piezoelectric plates   1343

TABLE I. Comparison between experimentally measured and theoretically calculated resonance frequencies for single side grating devices built on AT cut and Z cut plates of quartz (128 μm thick, grating's period 20 μm, strip height 150 nm).

Number of mode	Resonance frequency on Z cut plate of quartz in (MHz)		Resonance frequency on AT cut plate of quartz (in MHz)	
	Theory	Experiment	Theory	Experiment
1st	90.6	90.4	127.20	127.0
2nd	91.6	92.2	128.5	128.5
3rd	94.3	⋯	130.0	130.0
4th	98.9	97.8	132.4	132.5
5th	105.4	106.4	138.8	140.01
6th	113.6	117.2	143.1	144.0

pling than those on an AT-cut device by analogy with conventional BAW.[23] In this case, the simplified electrical boundary conditions of Sec. III may be no longer valid. However, even the largest piezoelectric coupling on quartz appears very weak when compared to highly coupled materials like LiNbO$_3$ or LiTaO$_3$. Deriving an accurate model of transverse plate modes on these substrates would require a complete re-write of rigorous electrical boundary conditions. It should be emphasized finally that the above mentioned discrepancies between our model and experiments do not exceed 3% of the corresponding frequency (worst case of the sixth mode) which is a satisfying result.

IV. CONCLUSION

A theoretical analysis of transverse waves propagating on piezoelectric plates of finite thickness supporting single or double side periodic grating has been proposed. It gives access to the physical characteristics of the wavelike mechanical and electrical amplitudes, dispersion relation and so on. Rigorous mechanical boundary conditions and simplified electrical boundary conditions have been used to compute the problem. This approach can be justified by the fact that all computations have been performed for quartz, which is a weakly coupled piezoelectric material allowing this kind of assumption. For strongly coupled materials like lithium tetraborate or GaPO$_4$, more rigorous electrical boundary conditions must be considered.

However, comparison with experimental devices built on AT-cut and Z-cut plates of quartz has shown the efficiency and the accuracy of the proposed model. Further experiments are now required to confirm some theoretical results concerning electrical boundary conditions or the influence of the plate thickness. Double side grating device properties also have to be checked. Other properties of these devices will be theoretically and experimentally investigated, particularly the gravimetric sensitivity of single side grating devices which could be applied for sensing applications.

APPENDIX: RELATIONS BETWEEN FIRST AND SECOND FLOQUET'S MODE

As defined in Sec. II A, the relation between β_0^{st} and β_0^{nd} (first and second Floquet's mode wave numbers) is written as follows:

$$\beta_0^{nd} = \frac{\pi}{\Lambda} - \beta_0^{st}, \tag{A1}$$

which provides, considering Eq. (2), the following relation for a given mode n:

$$\beta_n^{nd} = -\beta_{-(n+1)}^{st}; \tag{A2}$$

the relation between α_{nr}^{st} and α_{nr}^{nd} is provided by analyzing the solutions of Eq. (5). This equation can be written under the general following form:

$$\alpha_n^4 + \alpha_n^3 \beta_n P_3 + \alpha_n^2 \beta_n^2 P_2 + \alpha_n \beta_n^3 P_1 + \beta_n^4 P_0 = 0, \tag{A3}$$

where $P_3 \ldots {}_0$ are considered as constant complex coefficients. For a given value of $\beta_n = \beta_n^{st}$, Eq. (A3) provides four values α_{nr}^{st}. Replacing β_n by β_n^{nd} in (A3) and considering Eq. (A2) yield the following form of Eq. (A3):

$$\alpha_n^{nd4} - \alpha_n^{nd3} \beta_{-(n+1)}^{st} P_3 + \alpha_n^{nd2} \beta_{-(n+1)}^{st2} P_2$$
$$- \alpha_n^{nd} \beta_{-(n+1)}^{st3} P_1 + \beta_{-(n+1)}^{st4} P_0 = 0. \tag{A4}$$

Changing the sign of the odd powers of α_n in Eq. (A3) induces a change of the sign of the four complex solutions α_{nr}^{nd}. As a consequence, the four solutions α_{nr}^{nd} of Eq. (A4) can be related to the first Floquet's mode α_{nr}^{st} using the following relation:

$$\alpha_{nr}^{nd} = -\alpha_{-(n+1)r}^{st}. \tag{A5}$$

Equations (A2) and (A5) can now be used to express the relation between D_{nr}^{st} and D_{nr}^{nd} of which the basic expression is given by Eq. (7):

$$D_{nr}^{nd} = D_{-(n+1)r}^{st}. \tag{A6}$$

It is now necessary to compare boundary conditions of both the first and second Floquet's mode to establish relations between mechanical amplitudes A_{nr}^{st} and A_{nr}^{nd}.

Electrical boundary conditions of Eq. (17) are written for the second Floquet's mode considering Eqs. (A5) and (A6):

$$\sum_{r=1}^{4} D_{-(q+1)r}^{st} A_{qr}^{nd} e^{-\alpha_{-(q+1)r}^{st} e} = 0 \quad \text{along } a_2 = +e,$$

$$q = 0, \pm 1, \pm 2, \ldots, \tag{A7}$$

$$\sum_{r=1}^{4} D_{-(q+1)r}^{st} A_{qr}^{nd} e^{\alpha_{-(q+1)r}^{st} e} = 0 \quad \text{along } a_2 = -e.$$

Identically, Eqs. (23) are given as follows for the second Floquet's mode:

$$\sum_{r=1}^{4} (\alpha^{st}_{-(q+1)r} e_{26} + j\beta^{st}_{-(q+1)} e_{36}$$

$$+ D^{st}_{-(q+1)r} (\varepsilon_{22} \alpha^{st}_{-(q+1)r} + j\varepsilon_{23} \beta^{st}_{-(q+1)}$$

$$- |\beta^{st}_{-(q+1)}| \varepsilon_0 e^{-|\beta^{st}_{(q+1)}|e}))$$

$$\times A^{nd}_{qr} e^{-\alpha^{st}_{-(q+1)r}e} = 0 \quad \text{along} \quad a_2 = +e, \quad (A8)$$

$$\sum_{r=1}^{4} (\alpha^{st}_{-(q+1)r} e_{26} + j\beta^{st}_{-(q+1)} e_{36} + D^{st}_{-(q+1)r} (\varepsilon_{22} \alpha^{st}_{(q+1)r}$$

$$+ j\varepsilon_{23} \beta^{st}_{-(q+1)} + |\beta^{st}_{-(q+1)}| \varepsilon_0 e^{+|\beta^{st}_{(q+1)}|e}))$$

$$\times A^{nd}_{qr} e^{+\alpha^{st}_{-(q+1)r}e} = 0 \quad \text{along} \quad a_2 = -e.$$

Finally, the mechanical boundary conditions (16) written for the second Floquet's mode yields

$$\sum_{n=-\infty}^{+\infty} \sum_{r=1}^{4} \left[K^{st}_{-(q+1)-(n+1)r} - \frac{\Lambda}{h} \delta_{nq} (F^{st}_{-(n+1)r} \right.$$

$$\left. + D^{st}_{-(n+1)r} Z^{st}_{-(n+1)r}) \right] e^{-\alpha^{st}_{-(n+1)r}e} A^{nd}_{nr} = 0,$$

$$\sum_{n=-\infty}^{+\infty} \sum_{r=1}^{4} \left[L^{st}_{-(q+1)-(n+1)r} + \frac{\Lambda}{k} \delta_{nq} (F^{st}_{-(n+1)r} \right.$$ (A9)

$$\left. + D^{st}_{-(n+1)r} Z^{st}_{-(n+1)r}) \right] e^{\alpha^{st}_{-(n+1)r}e} A^{nd}_{nr} = 0.$$

From Eqs. (A7), (A8), and (A9), it can be established that calculating the second Floquet's mode is equivalent to calculating the first one but simultaneously changing the sign of a_2 and a_3. As a consequence, the amplitude of the second Floquet's mode can be related to A^{st}_{nr} only if the propagating structure is rigorously symmetric along a_2 and a_3.

This means that for a double side grating device, with h and k equal (no shift δ between the grating origins) and identical electric boundary conditions on the top and bottom surfaces, the first and second Floquet's mode amplitudes are related as follows:

$$A^{nd}_{nr} = A^{st}_{-(n+1)r}. \quad (A10)$$

Any difference between the top and bottom surfaces of the plate induces the invalidity of relation (A10), and no simple relation between A^{st}_{nr} and A^{nd}_{nr} can be deduced from boundary conditions of Eq. (16), (17), (23) and (A7), (A8), (A9).

It can even be noted that in classical STW propagation, because of unsymmetric boundary conditions, along a_2, A^{st}_{nr} and A^{nd}_{nr} are not simply connected. This can be very easily deduced from boundary conditions (16), (17), (23), (A7), (A8), and (A9) by setting $e > 35\lambda_a$. This procedure shows that the upper surface second Floquet's mode is equivalent to the lower surface first Floquet's mode, but does not allow us to establish any other relation between modes propagating on the same surface. This result shows that STW and more generally transverse waves trapped by surface gratings may exhibit some kind of directivity properties, which may be advantageously used if controlled.

[1] B. A. Auld, J. J. Gagnepain, and M. Tan, ''Horizontal shear waves on corrugated surfaces,'' Electron. Lett. **12**(24), 650–651 (1976).

[2] Yu. V. Gulyaev and V. P. Plesskii, Sov. Tech. Phys. Lett. **3**(3), 87–88 (1977).

[3] V. I. Anisimkin, M. A. Magomedov, and V. I. Fedorov, ''Resonators on surface transverse waves,'' Sov. Phys. Tech. Phys. Lett. **11**(1), 55–60 (1985).

[4] D. F. Thompson and B. A. Auld, ''Surface Transverse Wave propagation under metal strip gratings,'' Proc. 1986 Ultrasonics Symp., IEEE Cat. No. 86CH2375-4, pp. 261–266.

[5] A. Ronnekleiv, ''High Q resonators based on surface transverse waves,'' Proc. 1986 Ultrasonics Symp., IEEE Cat. No. 86CH2375-4, pp. 257–280.

[6] T. L. Bagwell and R. C. Bray, ''Novel surface transverse wave resonators with low loss and high Q,'' Proc. 1987 Ultrasonics Symp., IEEE Cat. No. 87CH2492-7, pp. 319–324.

[7] I. D. Avramov, ''1 GHz low loss coupled resonator filter using surface skimming bulk waves and Bleustein-Gulyaev waves,'' Electron. Lett. **5**(27), 414–415 (1991).

[8] R. Almar, B. Horine, and J. Andersen, ''High frequency STW resonator filters,'' Proc. 1992 Ultrasonics Symposium, IEEE Cat. No. 92CH3118-7, pp. 51–56.

[9] I. D. Avramov, ''Microwave oscillators stabilized with surface transverse wave resonant devices,'' proc. 46th Ann. Freq. Cont. Symp., Hershey, PA, June 1992.

[10] E. Bigler, B. A. Auld, E. Ritz, and E. Sang, ''An analysis of the influence of design parameters on the resonant frequency and Q-factor of surface transverse wave (STW) resonators,'' Proc. 45th Ann. Freq. Cont. Symp., Los Angeles, June 1991.

[11] S. Ballandras, E. Gavignet, and E. Bigler, ''Single and double side grating devices built on thin quartz plates for the excitation of surface transverse waves,'' 1995 IEEE Ultrasonics Symposium, Seattle, Nov. 1995.

[12] A. Renard, J. Henaff, and B. A. Auld, ''SH surface wave propagation on corrugated surfaces of rotated Y-cut quartz and berlinite crystals,'' 1981 Ultrasonics Symp. Proc. IEEE, Cat. No. 81CH1689-9, pp. 123–128.

[13] E. Danicki and D. Bogucki, ''Wave propagation and scattering in elastic plate with periodically grooved surface,'' Arch. Acoust. **18**(1), 113–130 (1993).

[14] E. Gavignet, S. Ballandras, and E. Bigler, ''Theoretical analysis of surface transverse waves propagating on a piezoelectric substrate under shallow-groove or thin metal strip grating,'' J. Appl. Phys. **77**(12), 6228–6233 (1995).

[15] E. Gavignet, S. Ballandras, and E. Bigler, ''Analysis and experimental study of surface transverse wave resonators on quartz,'' J. Appl. Phys. **79**(12), 8944–8950 (1996).

[16] IEEE Standard on Piezoelectricity 176-1949, Proc. IRE **37**, 1378–1395 (1949).

[17] A. J. Slobodnik, E. D. Conway, and R. T. Delmonico, *Microwave Acoustics Handbook, Vol. 1A: Surface Wave Velocities* (Air Force Cambridge Research Labs, 1973).

[18] M. A. Goodberlet and D. L. Lee, ''The excitation and detection of surface-generated bulk waves,'' IEEE Trans. Sonics Ultrason. **SU-31**(2), 67–76 (1984).

[19] J. Hou and H. Van de Vaart, ''Mass sensitivity of plate modes in surface acoustic wave devices and their potential as chemical sensors,'' Proc. 1987 Ultrasonics Symposium, IEEE Cat. No. 87CH2492-7, pp. 573–578.

[20] B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990), Vol. II, pp. 159–170.

[21] E. Bigler, E. Gavignet, B. A. Auld, E. Ritz, and E. Sang, ''Surface transverse wave (STW) quartz resonators in the GHz range,'' Proc. 6th European Frequency and Time Forum, Noordwijk, The Netherlands, March 1992, pp. 219–222.

[22] V. C. Plessky, ''A two parameter coupling of modes model for shear horizontal type SAW propagation in periodic gratings,'' Proc. 1993 Ultrasonics Symposium, IEEE Cat. No. 93CH3301-9, pp. 269–274.

[23] J. F. Rosenbaum, *Bulk Acoustic Wave Theory and Devices* (Artech House, 1988), pp. 400–410.

The elastodynamic response of a semi-infinite anisotropic solid to sudden surface loading

A. G. Every, K. Y. Kim,[a)] and A. A. Maznev[b)]
University of the Witwatersrand, PO WITS 2050, South Africa

Integral expressions are derived for the displacement response tensor of a semi-infinite anisotropic elastic continuum of unrestricted symmetry to a concentrated force suddenly applied to its surface. The surface response is reduced to a one-dimensional integral for numerical evaluation, while the interior response is left as a two-dimensional integral. Calculated surface response functions for Cu(001) are presented. These display multiple Rayleigh wave arrival singularities as well as bulk wave arrivals. Calculated interior response functions for Zn(0001) are presented. These display bulk and head wave arrivals. In followup papers these methods will be used in the interpretation of capillary fracture generated waveforms measured in a number of different materials. © *1997 Acoustical Society of America.* [S0001-4966(97)06208-5]

INTRODUCTION

There is growing interest in the dynamic response functions (Green's functions) of elastically anisotropic solids.[1–14] Part of the reason for this is the need for simple and efficient computational algorithms for use in interpreting waveform data emanating from laser ultrasound experiments,[15–18] transmission and reflection acoustic microscopy,[19–22] and various other transient wave experiments.[23,24] For some purposes it is sufficient to trace the progress of wave arrivals through the medium, making use of the ray approximation.[15–17] In some cases detailed transmission waveforms measured in samples of finite geometry can be reasonably well accounted for with response functions calculated for the infinite continuum.[1] Surfaces do, however, have a modifying effect on transmission waveforms, and this generally becomes more pronounced as the sensing point is moved away from epicenter and head waves come into prominence.[25] When the displacement is measured on the same surface as the applied force, the surface plays a determining role in the response, which tends to be dominated by one or more Rayleigh wave or pseudo-surface-wave arrivals.

This paper is the first in a three part series in which we show that surface and also to a large extent transmission waveforms, generated by a transient force acting on the surface of an elastically anisotropic solid, conform well to dynamic response functions calculated for the semi-infinite continuum. Reverberation effects arising from reflections from the opposite faces and sidewalls lie outside the scope of the present series of articles.

The problem of the displacement response of an elastic half-space to a point or line load suddenly applied at the surface has received wide attention over the years, since first being posed by Lamb.[26] Analytical solutions for point and line loading of an isotropic half-space were first obtained by Cagniard by a method involving Laplace transformation and intricate deformation of the integration contour to analytically perform one of the integrations in the inverse transformation. These celebrated solutions are to be found in a number of books on the dynamics of solids.[27,28] Burridge[29] extended the Cagniard method to calculate the response of an anisotropic half space to an impulsive line load at the surface, and similar results have been obtained through direct integration by Maznev and Every.[30] The Cagniard–de Hoop method has been extensively reviewed by van der Hijden.[31] Willis[32] in a seminal paper obtained the formal solutions to a wide class of self similar problems for the anisotropic half-space using Fourier and radon transforms, and this method has been further developed by Wang and Achenbach.[4] Payton[33] has treated a number of problems for transversely isotropic half-spaces that admit closed-form solutions. Recently Mourad *et al.*[5] have used the Cagniard–de Hoop method to calculate the interior response of anisotropic half spaces of hexagonal and cubic symmetry to point loading. Their method reduces ultimately to a single angular integral which has to be done numerically. Another approach that has been taken recently in calculating response functions of anisotropic half spaces is integral representation in terms of a δ function by Tewary and Fortunko.[13,14]

In this paper we establish methods for calculating the dynamic response at surface and interior points of a semi-infinite anisotropic elastic continuum of unrestricted symmetry to a point load suddenly applied at the surface. Our approach is to Fourier transform the equations of motion and boundary conditions with respect to time and the spatial coordinates parallel to the surface, solve the resulting algebraic equations, and then carry out the inverse transformation. The surface response is reduced to a one-dimensional integral for numerical evaluation, while the interior response is left as a two-dimensional integral. The method we use for computing these numerical integrals is able to cope with Rayleigh poles and pseudo-surface-wave resonances. Our method of calculating the interior response, in that it is in the form of a two dimensional integral, is computationally somewhat less efficient than the Cagniard–de Hoop method[5] and the method of

---

[a)]Present address: Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, NY 14853.
[b)]Present address: Department of Chemistry, MIT, Cambridge, MA 02139.

Wang and Achenbach,[4] but it has the advantage of conceptual simplicity and ease of implementation. In our analysis we give particular attention to wave arrival singularities, which are a striking feature of both surface and interior responses. By way of illustrative numerical example, we present calculated surface responses for Cu(001) and interior responses for Zn(0001). In subsequent papers in this series, we apply our method to the interpretation of capillary fracture generated surface and transmission waveforms in a number of different materials.

I. METHOD OF CALCULATION

We consider a general anisotropic elastic continuum of density ρ and elastic modulus tensor C_{ijkl} occupying the half-space $x_3 > 0$. A concentrated point force $\mathbf{F}(t) = (F_j \epsilon(t))$ with sign-function time dependence

$$\epsilon(t) = \begin{cases} 1, & t > 0 \\ -1, & t < 0 \end{cases} \tag{1}$$

acts at the origin on the otherwise free surface of the half-space. The displacement field $\mathbf{u}(\mathbf{x},t)$ in response to this force is given by

$$u_i(\mathbf{x},t) = G_{ij}^\epsilon(\mathbf{x},t) F_j, \tag{2}$$

where $G_{ij}^\epsilon(\mathbf{x},t)$ is the response function tensor.

Later in this paper we shift our attention to the response function $G_{ij}^\Theta(\mathbf{x},t)$ pertaining to a force with Heaviside step function time dependence

$$\Theta(t) = \tfrac{1}{2}\{\epsilon(t) + 1\}, \tag{3}$$

which is more amenable to direct measurement. Rather than calculating $G_{ij}^\Theta(\mathbf{x},t)$ directly, it is somewhat simpler to calculate $G_{ij}^\epsilon(\mathbf{x},t)$ and then obtain $G_{ij}^\Theta(\mathbf{x},t)$ from

$$G_{ij}^\Theta(\mathbf{x},t) = \tfrac{1}{2}\{G_{ij}^\epsilon(\mathbf{x},t) - G_{ij}^\epsilon(\mathbf{x},0)\}. \tag{4}$$

This sidesteps the δ function that occurs in the Fourier transform of $\Theta(t)$, but not that of $\epsilon(t)$. The time derivative of $G_{ij}^\Theta(\mathbf{x},t)$ is the dynamic Green's tensor of the medium.

The displacement field is required to satisfy the equations of motion

$$\rho \frac{\partial^2 u_i}{\partial t^2} = C_{ijkl} \frac{\partial^2 u_l}{\partial x_j \, \partial x_k}, \quad x_3 > 0, \tag{5}$$

subject to the boundary conditions on the stress tensor

$$\sigma_{\diagup 3}(\mathbf{x}_\parallel, x_3 = 0_+, t) = -F_j \delta_{jl} \delta(\mathbf{x}_\parallel) \epsilon(t), \tag{6}$$

where $\mathbf{x}_\parallel = (x_1, x_2)$ denotes the position vector in the surface and $\delta(\mathbf{x}_\parallel) = \delta(x_1)\delta(x_2)$ is the two-dimensional δ function. The negative sign in Eq. (6) arises from the fact that the surface tractions $\sigma_{j3}(\mathbf{x}_\parallel, x_3 = 0_+, t)$ are in reaction to the applied force. The initial conditions are $G_{ij}^\Theta(\mathbf{x},t) = 0$, $t < 0$.

Representing the boundary conditions in terms of their Fourier transform with respect to $\mathbf{x}_\parallel$ and t, we have

$$\sigma_{l3}(\mathbf{x}_\parallel, x_3 = 0_+, t) = \int_{-\infty}^{\infty} d^2 k_\parallel \int_{-\infty}^{\infty} d\omega \frac{F_j \delta_{jl}}{4\pi^3 i\omega}$$

$$\times \exp\{i(\mathbf{k}_\parallel \cdot \mathbf{x}_\parallel - \omega t)\}, \tag{7}$$

where $\mathbf{k}_\parallel = (k_1, k_2)$ is the projection of the wave vector in the surface and ω is the angular frequency.

We seek a solution to the equations of motion (5) and boundary conditions (7) in the form of a superposition of outgoing plane waves whose amplitudes are proportional to F_j:

$$u_i(\mathbf{x},t) = \int_{-\infty}^{\infty} d^2 k_\parallel \int_{-\infty}^{\infty} d\omega \sum_{n=1}^{3} A_j^{(n)} F_j U_i^{(n)}$$

$$\times \exp\{i(\mathbf{k}_\parallel \cdot \mathbf{x}_\parallel + k_3^{(n)} x_3 - \omega t)\}. \tag{8}$$

For each value of $\mathbf{k}_\parallel$ and ω, the third component k_3 of $\mathbf{k}$ and the polarization vector $\mathbf{U}$ are related by the Christoffel equations

$$(C_{ijkl} s_j s_k - \rho \delta_{il}) U_l = 0, \tag{9}$$

where $\mathbf{s} = \mathbf{k}/\omega$, is the acoustic slowness vector and $s_3^{(n)}$ are roots of the characteristic sextic equation

$$\det|C_{ijkl} s_j s_k - \rho \delta_{il}| = 0. \tag{10}$$

Equation (10) yields six solutions of which three are chosen which correspond to outgoing waves, on the basis that they are either homogeneous (bulk) waves with ray vectors directed into the interior, or inhomogeneous (evanescent) waves which decay into the interior.[34] The choice of outgoing waves depends on the sign of ω, which will be restricted to positive values, as explained later.

From the stress–strain relationship, $\sigma_{lm} = C_{lmpq} \partial u_p / \partial x_q$, and Eq. (8) it follows that the surface tractions are given by

$$\sigma_{l3}(\mathbf{x}_\parallel, x_3 = 0_+, t) = \int_{-\infty}^{\infty} d^2 k_\parallel \int_{-\infty}^{\infty} d\omega \, i\omega \sum_{n=1}^{3} A_j^{(n)} F_j B_l^{(n)}$$

$$\times \exp\{i(\mathbf{k}_\parallel \cdot \mathbf{x}_\parallel - \omega t)\}, \tag{11}$$

where

$$B_l^{(n)} = \sum_{pq} C_{3lpq} U_p^{(n)} s_q^{(n)}, \tag{12}$$

in which $(s_1^{(n)}, s_2^{(n)}) = (s_1, s_2)$. Comparing (7) and (11) we arrive at a set of three linear equations for the partial wave weighting factors $A_j^{(n)}$

$$\sum_{n=1}^{3} A_j^{(n)} B_l^{(n)} = -\delta_{jl}/4\pi^3 \omega^2. \tag{13}$$

These have solution

$$A_j^{(n)} = -\frac{1}{4\pi^3 \omega^2} (\mathbf{B}^{-1})_j^{(n)} = -\frac{1}{4\pi^3 \omega^2} \frac{\mathrm{adj}(\mathbf{B})_j^{(n)}}{\det|\mathbf{B}|}, \tag{14}$$

adj denoting matrix adjoint.

From Eqs. (2), (8), and (14) it follows that

$$G_{ij}^\epsilon(\mathbf{x},t) = -\frac{1}{4\pi^3} \int_{-\infty}^{\infty} d^2 k_\parallel \int_{-\infty}^{\infty} \frac{d\omega}{\omega^2} \sum_{n=1}^{3} \Lambda_{ij}^{(n)}$$

$$\times \exp\{i(\mathbf{k}_\parallel \cdot \mathbf{x}_\parallel + k_3^{(n)} x_3 - \omega t)\}, \tag{15}$$

where

$$\Lambda_{ij}^{(n)} = \frac{\text{adj}(\mathbf{B})_j^{(n)} U_i^{(n)}}{\det|\mathbf{B}|}. \tag{16}$$

Since $G_{ij}^{\epsilon}(\mathbf{x},t)$ is real, it follows that its temporal Fourier transform has the property $\tilde{G}_{ij}^{\epsilon*}(\mathbf{x},\omega) = \tilde{G}_{ij}^{\epsilon}(\mathbf{x},-\omega)$, where the asterisk denotes complex conjugation. We exploit this property to restrict the integration over ω in (15) to the interval $[0,\infty]$, retaining only the real part. On further replacing the integration variable $\mathbf{k}_{\parallel}$ by $\mathbf{s}_{\parallel} = \mathbf{k}_{\parallel}/\omega$ we arrive at the result

$$G_{ij}^{\epsilon}(\mathbf{x},t) = -\frac{1}{2\pi^3}\text{Re}\int_{-\infty}^{\infty} d^2 s_{\parallel} \int_0^{\infty} d\omega \sum_{n=1}^{3} \Lambda_{ij}^{(n)}$$

$$\times \exp\{i\omega(\mathbf{s}_{\parallel}\cdot\mathbf{x}_{\parallel} + s_3^{(n)} x_3 - t)\}. \tag{17}$$

We treat the surface response and the response at interior points separately below.

II. SURFACE RESPONSE

By choosing the coordinate system suitably one can, without loss of generality, take the observation point $\mathbf{x}$ to be along the x_1 axis on the surface. Setting $x_2 = x_3 = 0$, Eq. (17) simplifies to

$$G_{ij}^{\epsilon}(x_1,t) = -\frac{1}{2\pi^3}\text{Re}\int_{-\infty}^{\infty} ds_1 \int_0^{\infty} d\omega \Phi_{ij}(s_1)$$

$$\times \exp\{i\omega(s_1 x_1 - t)\}, \tag{18}$$

where

$$\Phi_{ij}(s_1) = \int_{-\infty}^{\infty} ds_2 \Psi_{ij}(\mathbf{s}_{\parallel}), \tag{19}$$

and

$$\Psi_{ij}(\mathbf{s}_{\parallel}) = \sum_{n=1}^{3} \Lambda_{ij}^{(n)}(\mathbf{s}_{\parallel}). \tag{20}$$

Convergence of the integral over ω is achieved by the replacement $(s_1 x_1 - t) \rightarrow (s_1 x_1 - t + i0_+)$. On integrating, one obtains

$$G_{ij}^{\epsilon}(x_1,t) = -\frac{1}{2\pi^3}\text{Re}\int_{-\infty}^{+\infty} ds_1 \frac{i\Phi_{ij}(s_1)}{(s_1 x_1 - t + i0_+)}$$

$$= -\frac{1}{2\pi^3}\left[-\mathscr{P}\int_{-\infty}^{+\infty} ds_1 \frac{\text{Im }\Phi_{ij}(s_1)}{(s_1 x_1 - t)}\right.$$

$$\left. + \frac{\pi}{|x_1|}\text{Re }\Phi_{ij}(t/x_1)\right], \tag{21}$$

with $\mathscr{P}\!\int$ denoting the principal value of the integral.

Using the fact that for negative times $G_{ij}^{\epsilon}(x_1,t)$ is constant, and setting $G_{ij}^{\epsilon}(-x_1,-t) = G_{ij}^{\epsilon}(-x_1,0)$, $t > 0$, we arrive at the Kramers–Kronig-type relation

$$-\mathscr{P}\int_{-\infty}^{+\infty} ds_1 \frac{\text{Im }\Phi_{ij}(s_1)}{s_1 x_1 - t} = \frac{\pi}{|x_1|}\text{Re }\Phi_{ij}(t/x_1)$$

$$+ 2\pi^3 G_{ij}^{\epsilon}(-x_1,0). \tag{22}$$

It follows immediately from Eqs. (4), (21), and (22) that the response function pertaining to a Heaviside force is given by

$$G_{ij}^{\Theta}(x_1,t>0) = -\frac{1}{2\pi^2|x_1|}\text{Re}\{\Phi_{ij}(t/x_1) - \Phi_{ij}(0)\}, \tag{23}$$

which, bearing in mind (19), is in the form of a one-dimensional integral. For an isotropic solid this integral can be evaluated analytically, but for an anisotropic solid, numerical methods are in general required for its evaluation.

A. Some properties of $\Psi_{ij}(\mathbf{s}_{\parallel})$, $\Phi_{ij}(s_1)$, and $G_{ij}^{\Theta}(x_1,t)$

For small values of $\mathbf{s}_{\parallel}$ lying in the fully supersonic region enclosed by the curve of critical longitudinal (L) slownesses in the surface, all $s_3^{(n)}(\mathbf{s}_{\parallel})$ are real, and hence $\Psi_{ij}(\mathbf{s}_{\parallel})$ is real. Outside this critical curve, some or all of the $s_3^{(n)} \times (\mathbf{s}_{\parallel})$ are complex or pure imaginary, and $\Psi_{ij}(\mathbf{s}_{\parallel})$ is in general complex.

Further restrictions are imposed by the presence of materials symmetry. In the case where the x_3 axis lies along a twofold axis or is perpendicular to a symmetry plane, the following conditions hold:

The individual components of Ψ_{ij} are either symmetric or antisymmetric with respect to reversal of the direction of $\mathbf{s}_{\parallel}$:

$$\Psi_{ij}(-\mathbf{s}_{\parallel}) = \Psi_{ij}(\mathbf{s}_{\parallel}), \quad ij \in (11,22,33,12,21),$$
$$\Psi_{ij}(-\mathbf{s}_{\parallel}) = -\Psi_{ij}(\mathbf{s}_{\parallel}), \quad ij \in (13,23,31,32). \tag{24}$$

Likewise, for reversal of the direction of x_1,

$$G_{ij}^{\Theta}(-x_1,t) = G_{ij}^{\Theta}(x_1,t), \quad ij \in (11,22,33,12,21),$$
$$G_{ij}^{\Theta}(-x_1,t) = -G_{ij}^{\Theta}(x_1,t), \quad ij \in (13,23,31,32). \tag{25}$$

For large $\mathbf{s}_{\parallel}$ lying in the subsonic region outside the curve of limiting slow transverse (ST) slownesses, or transonic state as it is called, the solutions of Eq. (10), which takes the form of a cubic in s_3^2 because of symmetry, are all of the form $\pm iq$ or $\pm(p \pm iq)$, p,q real. The three discarded solutions are consequently $-s_3^{(n)} = s_3^{(n)*}$, where $s_3^{(n)}$ are the three chosen solutions corresponding to outgoing waves. It follows that

$$\Psi_{ij}(\mathbf{s}_{\parallel},-s_3(\mathbf{s}_{\parallel})) = \Psi_{ij}^*(\mathbf{s}_{\parallel},s_3(\mathbf{s}_{\parallel})). \tag{26}$$

It is evident from Eqs. (9), (10), (12), (17), and (20) that

$$\Psi_{ij}(-\mathbf{s}_{\parallel},-s_3(\mathbf{s}_{\parallel})) = -\Psi_{ij}(\mathbf{s}_{\parallel},s_3(\mathbf{s}_{\parallel})), \tag{27}$$

at all points except where $B_i^{(n)}$ is singular. From Eqs. (24), (25), and (26) one thus infers that, in the subsonic region away from singular points,

$$\Psi_{ij}^*(\mathbf{s}_{\parallel}) = -\Psi_{ij}(\mathbf{s}_{\parallel}) \quad \text{and} \quad \Psi_{ij}(\mathbf{s}_{\parallel}) \text{ is pure imaginary,}$$
$$ij \in (11,22,33,12,21),$$
$$\Psi_{ij}^*(\mathbf{s}_{\parallel}) = \Psi_{ij}(\mathbf{s}_{\parallel}) \quad \text{and} \quad \Psi_{ij}(\mathbf{s}_{\parallel}) \text{ is real,} \tag{28}$$
$$ij \in (13,23,31,32).$$

In the subsonic region, for all except possibly a few isolated directions of $\mathbf{s}_{\parallel}$, there is a single Rayleigh surface wave (RW) corresponding to the vanishing of $\det|B(\mathbf{s}_{\parallel})|$ (see
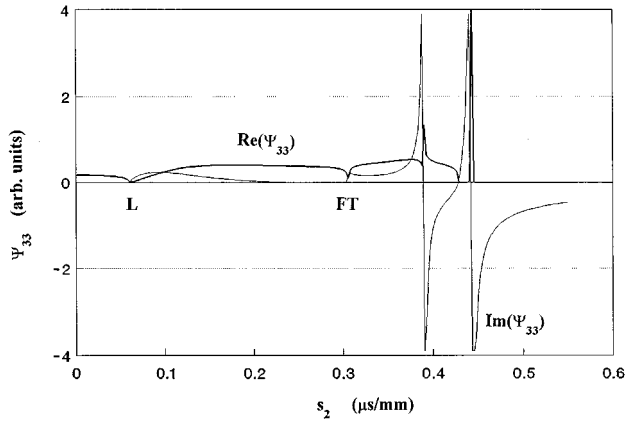
FIG. 1. Ψ_{33} ($s_1 = 0.21\,\mu s/mm$, s_2) for the (001) surface of copper. Constants for the calculation are $C_{11} = 169.0$, $C_{12} = 122.0$, and $C_{44} = 75.3$ in GPa, and $\rho = 8960$ kg/m^3.

Ref. 35). This gives rise, for a given direction, to a simple pole in $\Psi_{ij}(s_\parallel)$ at s_R, which on being moved above the real axis yields, for $s_\parallel$ near s_R,

$$\Psi_{ij}(s_\parallel) \sim \text{Res}_{s=s_R} \Psi_{ij}(s_\parallel) \left\{ \mathscr{P}\left(\frac{1}{s_\parallel - s_R}\right) + i\pi\delta(s_\parallel - s_R) \right\}, \tag{29}$$

where $\text{Res}_{s=s_R} \Psi_{ij}(s_\parallel)$ denotes the residue of $\Psi_{ij}(s_\parallel)$ at $s_\parallel = s_R$. Hence for $s_\parallel$ near s_R,

$$\text{Re}(\Psi_{ij}(s_\parallel)) \sim i\pi[\text{Res}_{s=s_R}\Psi_{ij}(s_\parallel)]\delta(s_\parallel - s_R),$$

$$ij \in (11,22,33,12,21),$$

$$\text{Re}(\Psi_{ij}(s_\parallel)) \sim [\text{Res}_{s=s_R}\Psi_{ij}(s_\parallel)]\mathscr{P}\left(\frac{1}{s_\parallel - s_R}\right), \tag{30}$$

$$ij \in (13,23,31,32).$$

For large s_1, beyond where the line $s_1 = $ constant intersects the RW slowness curve,

$$\text{Re}(\Psi_{ij}(\mathbf{s}_\parallel)) = 0, \quad (\Phi_{ij}(s_1)) = 0$$

and

$$\tag{31}$$

$$G_{ij}^{\Theta}(x_1,t) = \frac{\Phi_{ij}(0)}{2\pi^2|x_1|} = \text{const}$$

$$\text{for } ij \in (11,22,33,12,21).$$

If, further, the (x_1,x_3) plane is a materials symmetry plane, then

$$G_{ij}^{\Theta}(x_1,t) \equiv 0 \quad \text{for } ij \in (32,23,21,12). \tag{32}$$

B. Numerical integration and results

For an anisotropic solid the integration over s_2 in (19) to obtain Φ_{ij} has, in general, to be done numerically, and the method of integration has to cope with the RW poles and, in cases where they exist, pseudo-surface-acoustic-wave (PSAW) resonances.[34] This point is brought out in Fig. 1, which shows the dependence of Ψ_{33} on s_2 for fixed $s_1 = 0.21\,\mu s/mm$ for the (001) surface of copper. The reference
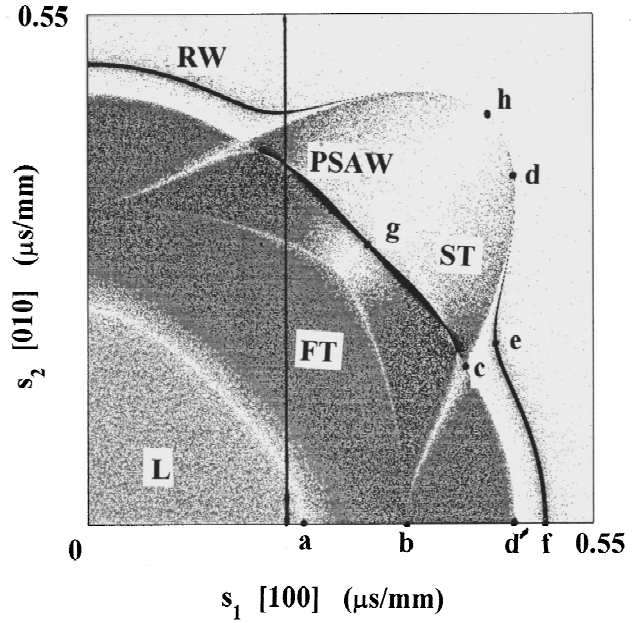
FIG. 2. Gray-scale representation of $\text{Re}(\Psi_{33}(s_1,s_2))$ for Cu(001).

axes are aligned along the principal crystallographic directions, and Ψ_{33} is unchanged by a reversal in the sign of s_1 or s_2. The real part of Ψ_{33} displays a sharp peak at 0.39 $\mu s/mm$, which is associated with a PSAW resonance. For $s_2 > 0.43$ $\mu s/mm$, Ψ_{33} is pure imaginary except for a δ-function singularity at 0.45 $\mu s/mm$, associated with the RW pole. Sharp dips occur in $\text{Re}(\Psi_{33})$ at the limiting L and one of the limiting fast transverse (FT) slownesses.

A global picture is provided by Fig. 2, which portrays the dependence of $\text{Re}(\Psi_{33})$ on both s_1 and s_2 as a gray-scale "image," with degree of darkness corresponding to the magnitude of $\text{Re}(\Psi_{33})$. To render the RW singularity visible, it has been artificially broadened by making the replacement

$$\Psi_{33} \to 1/(1/\Psi_{33} + \alpha), \tag{33}$$

where α is a small but finite positive number. A similar effect can be achieved by ascribing a small imaginary part to $\mathbf{s}_\parallel$. The vertical line in this image represents the "line scan" for the data in Fig. 1. The continuously shaded area in Fig. 2 comprises the Ψ_{33} weighted projections $\mathbf{s}_\parallel$ of the slowness vectors of all bulk modes in the first quadrant, its outer boundary being the so-called transonic state. The lines which stand out as lighter and which partition this domain, lie on the locus of limiting slownesses. This locus and the transonic state are the projection on the (s_1,s_2) plane of points on the three sheets of the slowness surface where the surface normal is parallel to the (s_1,s_2) plane. Along the [110] direction, the diagonal in Fig. 2, the ST bulk modes are SH polarized normal to the $(\bar{1}10)$ symmetry plane, and thus do not contribute to $\text{Re}(\Psi_{33})$, as is evident from the fading of the ST modes in this region. The sharp RW resonance is most intense in the $\langle 100 \rangle$ directions, and fades away towards the [110] direction, where it degenerates with the transonic state.

In the region extending about 20° to either side of the [110] direction there is a PSAW resonance lying within the band of ST bulk modes. Exactly in the [110] direction where

the RW degenerates, the PSAW becomes a pure supersonic two component surface acoustic wave (SAW).

The method we have used in calculating $G_{ij}^{\Theta}(x_1,t)$, is to make the replacement (33) and then evaluate the integral (19) numerically for 250 values of t. This is consistent in spirit with the numerical approach to inversion taken by a number of authors (see Ref. 7 and papers cited therein). Over most of the range of s_2, Ψ_{ij} is slowly varying, but in isolated regions there are the sharp RW and PSAW resonances, and kinks at limiting branch slownesses to contend with. We have dealt with this problem by dividing the range of s_2 into a number of intervals, a thousand is more than sufficient, and applying Rhomberg integration to each interval. It is only the few intervals where Ψ_{ij} is rapidly varying that more than one Rhomberg iteration is required to achieve convergence.

The above method has the advantage of simplicity and generality. With the coding we have implemented, it takes in the region of 15 min on a 133-MHz pentium personal computer to generate surface response functions of the type displayed in this paper, but this time can be considerably reduced at the cost of rounding of the sharp features. For instance, the response function for the fiber composite in Fig. 1 of Tewary and Fortunko[14] can be reproduced in less than one minute, or less still if fewer than 250 values of t are taken. Bearing in mind that our coding has been set up for general anisotropy, whereas that of Ref. 14 is specialized to tetragonal symmetry, it appears that their and our algorithms are comparable in numerical efficiency.

C. Wave arrival singularities

Figure 3(a) shows the surface response function $G_{33}^{\Theta}(x_1=100$ mm$,x_2=0,t)$ calculated for the (001) surface of copper. The response is zero until $t_a=23.1$ μs, at which moment there is a sudden downward kink corresponding to a discontinuous change in velocity. This event is associated with the sharp dip in Re(Ψ_{33}) at the limiting longitudinal slowness labeled a in Fig. 2, which is a branch point in the complex function Ψ_{33}. The kinks in G_{33}^{Θ} at $t_b=34.4$ μs and $t_c=40.2$ μs correspond to the limiting transverse slownesses labeled b and c in Fig. 2. These are bulk wave arrival singularities, and correspond to points on the slowness surface where the normal points in the observation direction. These singularities propagate outwards from the point of excitation at the bulk wave group velocities in that direction. Bulk wave singularities and the analytic form they take in the response functions of infinite anisotropic continua have been surveyed in Ref. 1. It is evident that a surface response function must also show nonanalytic behavior as the integration path $s_1=$const makes tangential contact with a limiting slowness curve, since to one side it has a double intersection with this curve, which is a locus of branch points, while on the other side it has no intersection.

At the surface there are also RW wave arrival singularities. These are generally much more prominent than the bulk wave arrivals, and so we devote more attention to them here. RW arrivals are conditioned by tangency of the line $s_1=s_R=t_R/|x_1|$ to the RW slowness curve. At the point of tangency the RW group velocity, which is normal to the RW

(a)

(b)

(c)
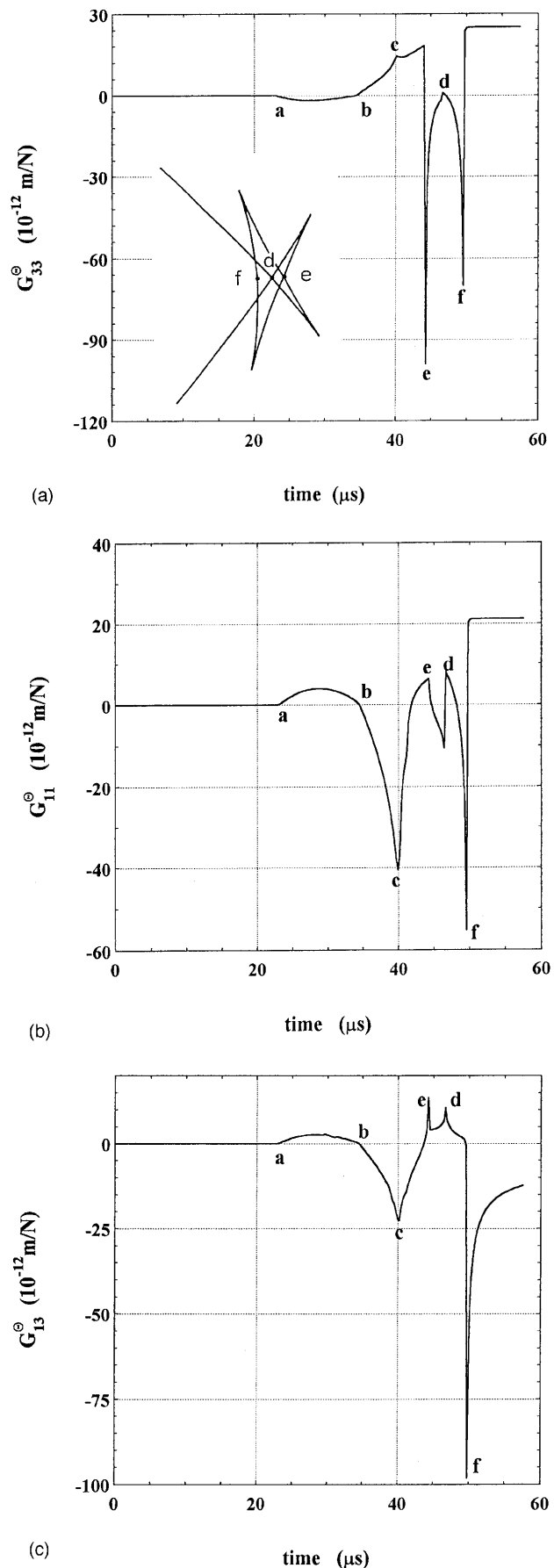
FIG. 3. Surface responses G_{ij}^{Θ} ($x_1=100$ mm, $x_2=x_3=0,t$) for Cu(001). (a) G_{33}^{Θ}, (b) G_{11}^{Θ}, (c) G_{13}^{Θ}. The insert in (a) shows a portion of the RW group velocity curve near the $\langle 100 \rangle$ direction.

slowness curve and hence parallel to the x_1 axis, is of magnitude $1/s_R$, and this is the velocity with which the wave arrival singularity propagates along the surface. The analytical form of G_{ij}^{Θ} for t near t_R depends on whether the surface of the half-space is a materials symmetry plane or not, and if it is, whether $ij \in (11,22,33,12,21)$ or $ij \in (13,23,31,32)$, and finally whether the RW slowness curve is convex or concave at the point of tangency.

Taking the surface to be a materials symmetry plane, $ij \in (11,22,33,12,21)$, and the RW slowness curve to be convex, then for $t < t_R$ the line $s_1 = t/|x_1|$ intersects the RW slowness curve twice near the point of tangency. In this region, the RW slowness is, in the parabolic approximation with the origin for s_2 located at the point of tangency, given by $s_1 = s_R - \beta s_2^2$, where β is the (positive) curvature. From Eqs. (19), (23), and (30) the singular part of G_{ij}^{Θ} thus takes the form

$$G_{ij}^{\Theta}(x_1,t) = \frac{A}{2\pi|x_1|} \int_{-\infty}^{\infty} ds_2 \ \delta\left(\beta s_2^2 - \left[\frac{t_R - t}{|x_1|}\right]\right), \qquad (34)$$

where $A = -i \ \mathrm{Res}_{s=s_R} \Psi_{ij}$ is a real constant for the purpose of the integration (although it does depend on the direction of x_1). The integral is readily evaluated, yielding

$$G_{ij}^{\Theta}(x_1,t) = \frac{A}{2\pi|x_1|^{1/2}\beta^{1/2}(t_R - t)^{1/2}}. \qquad (35)$$

For $t > t_R$, $\Psi_{ij} = 0$ in the region of tangency, and so the integral vanishes and G_{ij}^{Θ} is given by the constant term in Eq. (23). The wave arrival is thus led by an inverse square-root divergence and terminated by an infinite discontinuity. It may happen that there are other intersections of the line $s = s_R$ with the closed RW slowness curve, in which case there is a background in the neighborhood of t_R, which varies linearly with t, but after the last intersection, G_{ij}^{Θ} is exactly constant.

If at the point of tangency the RW slowness curve is concave, then the wave arrival is initiated by an infinite discontinuity, after which G_{ij}^{Θ} varies as $(t - t_R)^{-1/2}$.

For $ij \in (13,23,31,32)$ and the RW slowness curve concave, reference to Eqs. (19), (23), and (30) shows that the singular part of G_{ij}^{Θ} takes the form

$$G_{ij}^{\Theta}(x_1,t) = \frac{A}{2\pi^2|x_1|} \mathscr{P} \int_{-\infty}^{\infty} \frac{ds_2}{\left(\beta s_2^2 + \left[\frac{t_R - t}{|x_1|}\right]\right)}. \qquad (36)$$

On integrating, for $t < t_R$ this yields Eq. (35). For $t > t_R$ the path of integration intersects the RW slowness curve twice, and on taking the principal part of the integral, the result is zero, so that G_{ij}^{Θ} is constant. When the RW slowness curve is convex, these conditions are reversed, and the divergence follows the RW arrival.

In summary:

$$G_{ij}^{\Theta}(x_1,t) = \begin{cases} \dfrac{A_<}{2\pi|x_1|^{1/2}\beta^{1/2}(t_R - t)^{1/2}}, & t < t_R \\[12pt] \dfrac{A_>}{2\pi^2|x_1|}, & t > t_R, \end{cases} \qquad (37)$$

for $ij \in (11,22,33,12,21)$,   RW slowness curve convex,

   $ij \in (13,23,31,32)$,   RW slowness curve convcave,

where $A_>$ and $A_<$ are constants, and

$$G_{ij}^{\Theta}(x_1,t) = \begin{cases} \dfrac{A_<}{2\pi^2|x_1|}, & t < t_R \\[12pt] \dfrac{A_>}{2\pi|x_1|^{1/2}\beta^{1/2}(t - t_R)^{1/2}}, & t > t_R, \end{cases} \qquad (38)$$

for $ij \in (11,22,33,12,21)$,   RW slowness curve concave,

   $ij \in (13,23,31,32)$,   RW slowness curve convex.

The singular behavior displayed by the analytical solutions for isotropic solids[27,28] conforms to the above rules.

In the situation where the surface is not a materials symmetry plane, $\mathrm{Re}\ s_{s=s_R}\Psi_{ij}$ has both real and imaginary parts for all ij, and it follows that G_{ij}^{Θ} diverges on both sides of the RW arrival in accordance with

$$G_{ij}^{\Theta}(x_1,t) = \frac{A_{\gtrless}}{2\pi|x_1|^{1/2}|\beta|^{1/2}|t_R - t|^{1/2}}. \qquad (39)$$

The above results are consistent with Willis'[32] treatment of RW arrival singularities.

In anisotropic solids the displacement leading or trailing the RW arrival can vary enormously with direction. One source of this variation is the factor A. In directions for which the plane of polarization of the RW is perpendicular to the i or j axis, A is zero. For all ij, A also vanishes at points where the RW degenerates with the transonic state, as is evident in Fig. 2.

A second source of variation is that the singular part of the waveform is proportional to $|\beta|^{-1/2}$, and hence the wave intensity is proportional to $|\beta|^{-1}$. Variation in intensity through the curvature of the RW slowness curve in this way is known as surface phonon focusing, and has been studied theoretically by a number of authors.[36–39] The effects have been observed experimentally with laser generated ultrasonic SAW[18] and ballistic surface phonons.[40]

At points of inflection where $\beta = 0$, the parabolic approximation for the RW slowness equation breaks down. Taking x_1 and s_1 in the direction of the normal to the RW slowness curve at the point of inflection, the RW slowness equation to leading order now takes the form $s_1 = s_R - \gamma s_2^3$, $\gamma = $ constant. On integrating over s_2, one readily establishes that the singular part of G_{ij}^{Θ} near the wave arrival has the form

$$G_{ij}^{\Theta}(x_1,t) = \frac{A_{\gtrless}}{6\pi|x_1|^{1/3}|\gamma|^{1/3}|t_R - t|^{2/3}}. \qquad (40)$$

The inverse square-root dependence on x_1 and β for normal points, and the inverse third power dependence on x_1 and γ for points of inflection, is a characteristic also of frequency domain surface Green's functions in the asymptotic far-field limit.[36,38,39]

If the direction of the x_1 axis is changed slightly, then for one sense of rotation there are now two values of s_1 at which tangency occurs, and thus two wave arrivals propagating at slightly different group velocities. For the opposite sense of rotation there are no longer any points of tangency near the inflection point, and so no wave arrivals. It follows thereby that the point of inflection maps onto a cusp in the group velocity curve. Criteria for the existence of cusps in respect of the (100) and (111) surfaces of cubic crystals have been established by Maznev and Every.[41] Extending some distance beyond the cusp, where there is no longer any wave arrival in the strict sense, is a quasisingular feature which we refer to as an eidolon.[42] It has been observed for laser generated SAW by Maznev *et al.*[43]

Referring again to Fig. 3(a), the sharp dip at $t_e = 44.3$ μs is a RW singularity having the analytic form (38). The curvature of the RW slowness curve near point e is negative in this case, and so the RW arrival is manifested as a sudden dip followed by an initially steep continuous rise. The sharp dip at $t_f = 49.5$ μs is a RW singularity having the analytic form (37). It is shaped the familiar way around as encountered in the response function G^{Θ}_{33} of isotropic solids, because in this case the RW slowness curve near f is convex. Beyond t_f, G^{Θ}_{33} is constant, because $\text{Re}(\Psi_{33})$ is zero. There is a third RW singularity of the form (37) at t_d, but it shows up very faintly in the calculated response because, as can be seen in Fig. 2, point d is very close to where the RW degenerates with the transonic state and becomes SH in character. It also occurs very close to the transonic state d', and it is difficult to distinguish the two features from each other in this plot. The inset in Fig. 3(a) shows the folded character of the RW group velocity curve for Cu(001) near the x_1 axis. The three aforementioned wave arrivals correspond to the three points where the x_1 axis intersects this curve.

Figure 4 is a gray-scale representation of the dependence of $\text{Re}(\Psi_{11})$ on s_1 and s_2. Since Ψ_{11} selects out the s_1 axis, this plot does not display mirror symmetry across the diagonal (reflecting Ψ_{11} across the diagonal in fact yields Ψ_{22}). In the region of the s_1 axis $\text{Re}(\Psi_{11})$ shows a prominent peak at the L threshold (point a), rather than the dip as in the case of $\text{Re}(\Psi_{33})$. This quasisurface L mode resonance has been reported on before, and shows up in phonon imaging,[44–46] surface Brillouin scattering[47,48] and elsewhere. Towards the s_2 axis this resonance disappears (point b) because the L mode is polarized normal to the s_1 axis and so does not contribute to $\text{Re}(\Psi_{11})$. Here the L mode resonance is at a maximum in $\text{Re}(\Psi_{22})$. The RW, which is composed of inhomogeneous waves of predominantly L and T character, shows similar behavior to the L resonance, being most intense near the s_1 axis and fading to zero towards the s_2 axis.

Figure 3(b) shows the surface response function $G^{\Theta}_{11}(x_1 = 100 \text{ mm}, x_2, t)$ for Cu(001). As can be seen, the L bulk wave arrival is much more pronounced than for G^{Θ}_{33}, and the same applies to the bulk wave feature c, which can
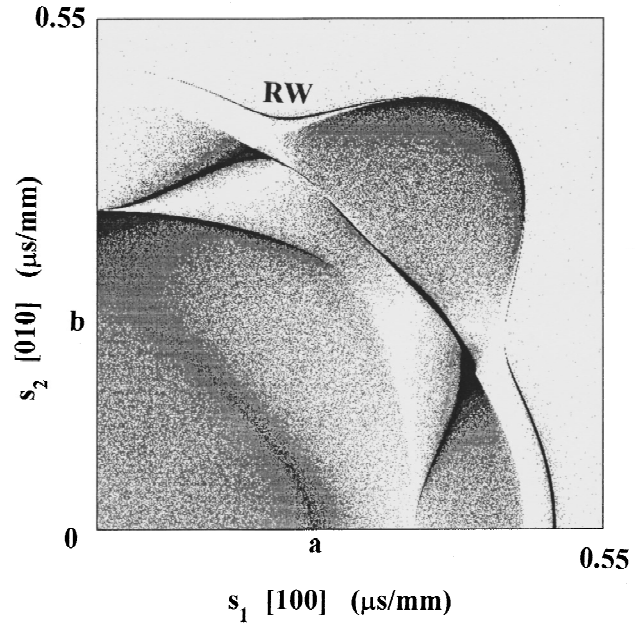
FIG. 4. Gray-scale representation of $\text{Re}(\Psi_{11}(s_1, s_2))$ for Cu(001).

be understood by reference to Fig. 4. The RW singularities are the same way around as for G^{Θ}_{33}, except that in this case feature d is more pronounced and e much less so, which can be understood by reference to Fig. 4.

Figure 3(c) shows the response function $G^{\Theta}_{13}(x_1 = 100 \text{ mm}, x_2, t)$ for Cu(001). One notes that in this case it is the wave arrival e that is of type (37), and the wave arrival f that is of type (38), so that after this last wave arrival there is a smooth asymptotic approach to the static response. Because the (x_1, x_3) plane is a crystallographic symmetry plane for Cu(001), $G^{\Theta}_{ij}(x_1, t) = 0$ for $ij \in 12,21,32,23$.

III. INTERIOR RESPONSE

Without loss of generality, we locate the observation point in the (x_1, x_3) plane. In the case of homogeneous waves corresponding to real $s_3^{(n)}$, convergence of the integral over ω in Eq. (16) is achieved through the replacement $(s_1 x_1 + s_3^{(n)} x_3 - t) \rightarrow (s_1 x_1 + s_3^{(n)} x_3 - t + i0_+)$. For inhomogeneous waves, $s_3^{(n)} x_3$ already comprises a positive imaginary part, and convergence is assured without the abovementioned replacement. On integrating over ω we obtain

$$
G^{\epsilon}_{ij}(\mathbf{x}, t) = -\frac{1}{2\pi^3 |\mathbf{x}|} \text{Re} \int_{-\infty}^{\infty} d^2 s_{\parallel}
$$
$$
\times \sum_{n=1}^{3} \frac{i\Lambda_{ij}^{(n)}(\mathbf{s}_{\parallel})}{s_1 \sin\theta + s_3^{(n)} \cos\theta - t/|\mathbf{x}| + i0_+}, \quad (41)
$$

where $\tan\theta = x_1 / x_3$. Essentially the same result is to be found in Willis[32] and in Wang and Achenbach,[4] except that they transform to radial and angular coordinates for $\mathbf{s}_{\parallel}$.

A. Numerical integration

To obtain the full time dependence of $G_{ij}(\mathbf{x}, t)$ at a given point $\mathbf{x}$, $s_3^{(n)}$ and $\Lambda_{ij}^{(n)}$ have to be determined for a two-dimensional (2D) manifold of points $\mathbf{s}_{\parallel}$. Discretization

for numerical purposes reduces this to calculating $s_3^{(n)}$ and $\Lambda_{ij}^{(n)}$ at say $\sim 10^3 \times 10^3 = 10^6$ points $\mathbf{s}_{\parallel}$, depending on the required accuracy. Treating (41) in a simpleminded way as a summation to obtain $G_{ij}(\mathbf{x},t)$ for say 240 values of t, involves an additional $240 \times 3 \times 10^6 \approx 10^9$ calculations, which in our experience, takes almost twice as long as the calculating of $s_3^{(n)}$ and $\Lambda_{ij}^{(n)}$. This is an unacceptable computational overhead. This problem has long been recognized, and sophisticated methods have been proposed for overcoming it. The Cagniard–de Hoop[5] method and the method developed by Willis[32] and Wang and Achenbach,[4] all reduce the evaluation of $G_{ij}(\mathbf{x},t)$ to a one-dimensional (1-D) angular integral, in which the integrand has to be determined iteratively for each value of t. The latter authors accomplish this by changing to radial and angular coordinates for $\mathbf{s}_{\parallel}$ and evaluating the integral over $|\mathbf{s}_{\parallel}|$ using the Cauchy residue theorem. This requires finding all the zeros in the complex plane of the denominator in (41) and determining the derivative of this denominator with respect to $|\mathbf{s}_{\parallel}|$. Information about $\Lambda_{ij}^{(n)}$ over the 2-D manifold is still required, but the overhead is considerably reduced. The Cagniard–de Hoop method has been widely applied to isotropic and transversely isotropic solids,[33] but only recently has it been implemented computationally to a significant extent for anisotropic solids.[5]

We have adopted the following numerical scheme for calculating $G_{ij}(\mathbf{x},t)$, which is much simpler than the above-mentioned methods, but which nevertheless, like them, substantially reduces the computational overhead. We proceed by casting (41) in the form

$$G_{ij}^{\epsilon}(\mathbf{x},t) = -\frac{1}{2\pi^3 |\mathbf{x}|} \mathrm{Re} \left\{ \int_{-\infty}^{\infty} \frac{f_{ij}(u)du}{u - t/|\mathbf{x}| + i0_+} \right.$$
$$\left. + \int_{-\infty}^{\infty} du \int_0^{\infty} \frac{g_{ij}(u,v)dv}{u + iv - t/|\mathbf{x}|} \right\}, \quad (42)$$

where the auxiliary functions $f_{ij}(u)$ and $g_{ij}(u,v)$ are given by

$$f_{ij}(u) = \sum_{n=1}^{3} \int_{\Omega_n} d^2 s_{\parallel} \delta(u - \mathrm{Re}[s_1 \sin\theta + s_3^{(n)} \cos\theta]) i\Lambda_{ij}^{(n)}, \quad (43)$$

$$g_{ij}(u,v) = \sum_{n=1}^{3} \int_{\Omega_n'} d^2 s_{\parallel} \delta(u - \mathrm{Re}[s_1 \sin\theta + s_3^{(n)} \cos\theta]) \delta(v - \mathrm{Im}[s_3^{(n)} \cos\theta]) i\Lambda_{ij}^{(n)}, \quad (44)$$

where Ω_n and Ω_n' are the domains in the (s_1,s_2) plane in which $s_3^{(n)}$ is real and complex, respectively.

All the detailed structure of $G_{ij}(\mathbf{x},t)$ is derived from $f_{ij}(u)$, while $g_{ij}(u,v)$, which tends linearly with v to zero towards the axis $v=0$, provides merely a smooth continuous background. For numerical purposes we divide the line u into a large number (7680) of intervals, and generate $f_{ij}(u)$ as a histogram of the accumulated values of $i\Lambda_{ij}^{(n)}$ for each interval in accordance with (43). The 1-D integral in (42), evaluated for 240 values of t thereby involves $\sim 1.8 \times 10^6$

calculations. In performing the 2-D integral, we divide the upper half of the (u,v) plane into a course grid of 240 $\times 120$ cells, and generate a 2-D histogram in accordance with (44), and then carry out the sum for the 240 values of t, which involves $\sim 6.9 \times 10^6$ calculations. The overhead is reduced in this way by more than a factor of 10 to less than 20% of the total computational time, which is acceptable. The computational time for the interior response functions presented below is under 15 min, which is less than that required to generate the surface response functions provided earlier, which nominally involve only 1-D numerical integration. The coding was kept general, with no advantage taken of symmetry other than taking the sample surface to be a materials symmetry plane to render Eq. (10) cubic in s_3^2.

It is appropriate at this point to comment on the relative merits and disadvantages of the contour integration methods mentioned earlier, as compared with the numerical approach described above. One could argue that, as a matter of principle, it is better to do as many integrations analytically as possible and only resort to numerical integration as a last resort. It might also be argued that integrals over a finite domain, as in Wang and Achenbach,[4] are more amenable to evaluation than the infinite integrals in our approach. Successful implementation of the method of Wang and Achenbach for a point force would, however, require the use of a robust and efficient algorithm for locating the poles in the complex plane, and for keeping track of the branch points and Riemann cuts. These calculations would have to be embedded in loops for varying t and the angular coordinate ϕ of $\mathbf{s}$. Divergences of the integrand in the integration over ϕ would have to be accommodated. We would anticipate that it is nevertheless possible to achieve greater numerical efficiency in this way, but at the cost of increased complexity in the coding. In progressing onto more complicated systems such as joined anisotropic half spaces and layered solids the problems will be more acute, and there may be a clear advantage to adopting a simple numerical approach of the type we have introduced here.

B. Results

Figure 5(a) shows calculated half-space response functions G_{33}^{Θ} for hexagonal zinc, with the surface of the half-space normal to the principal crystallographic axis. Three observation points have been chosen, namely $x_1 = 0$, 5, and 10 mm, with $x_3 = 25.8$ mm, partly for comparison with published experimental data.[7] These display a number of bulk wave arrival singularities labeled L (a discontinuity), I (a logarithmic discontinuity), S (a discontinuity followed by a gradual rise and leveling off), X, and C. These wave arrivals correspond to intersections of the viewing direction with sheets of the wave surface of zinc as shown in Fig. 6. In the epicentral direction, $x_1 = 0$, the two quasi-T sheets I and S of the wave surface intersect at a conical point, and the response shows a negative square-root divergence, which is associated with the phenomenon of external conical refraction.[23,24] The intersection F leads to a kink and a discontinuity which are barely perceptible in the response. For $x_1 = 10$ mm the I and F wave arrivals coincide at C, at a

FIG. 6. Zonal section of the qL and qT sheets of the wave surface of zinc.
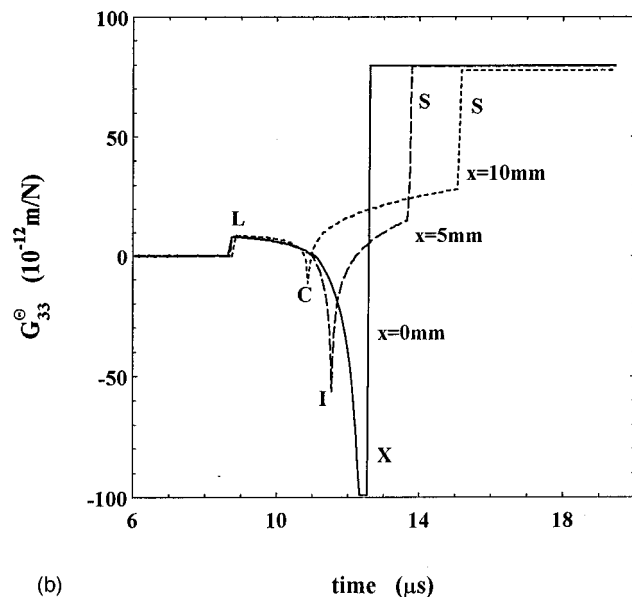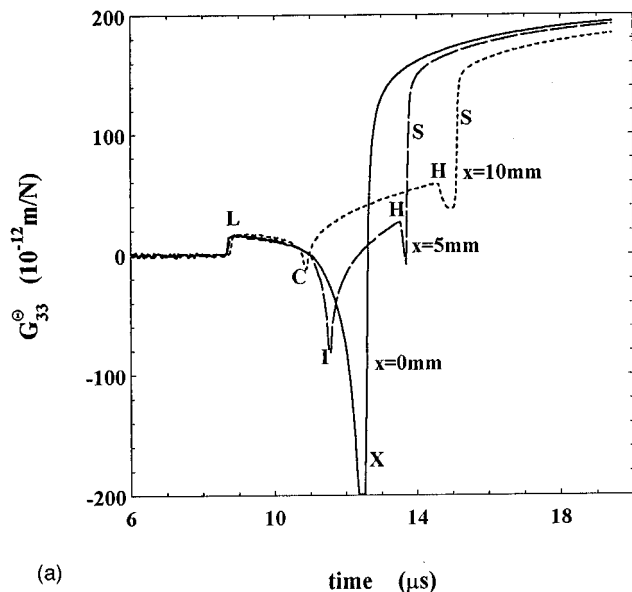
FIG. 5. (a) G_{33}^{Θ} ($x_1 = x, x_2 = 0, x_3 = 25.8$ mm) for a (0001)-oriented zinc semi-infinite continuum at interior points $x = 0$ (epicenter), 5 mm, and 10 mm. Constants for the calculation are $C_{11} = 163.75, C_{12} = 36.28, C_{13} = 52.48, C_{33} = 62.93$, and $C_{44} = 38.68$ in GPa, and $\rho = 7140$ kg/m³. (b) G_{33}^{Θ} for the same points in a zinc infinite continuum.

reduced in magnitude by about a factor of 2 as compared with those for the half-space, as one would expect. The bulk wave arrivals are of the same analytic form and occur at the same times as for the half-space. Head wave effects are of course absent from the infinite continuum results. Another difference is the constant displacement in the infinite continuum response following the last S arrival, as compared with the asymptotic leveling off in the half-space response.

IV. CONCLUSIONS

We have derived integral expressions for the surface and interior displacement response of a semi-infinite anisotropic elastic continuum subjected to sudden surface loading. By way of numerical example, calculated surface responses of Cu(001) and interior responses of Zn(0001) have been presented. The surface responses display bulk and surface wave arrival singularities, and the interior responses display bulk and head wave arrival singularities. In follow up papers we will use these methods in the interpretation of capillary fracture generated waveforms measured in a number of different materials.

cuspidal edge of the wave surface. There are no prominent singular features associated with the pure-T sheet of the wave surface, which is therefore not shown in Fig. 6. Slightly preceding the S arrival is the head wave arrival H, whose geometrical relation to the wave surface is shown in Fig. 6. Zinc is somewhat unusual in that the head wave merges with the wave surface at a point very close to the conical point.[25] These calculated responses are in very good agreement with waveforms measured in zinc by Kim et al.,[49,50] particularly with regard to the bulk and head wave arrivals.

Figure 5(b) shows zinc infinite continuum response functions calculated for the same configurations as above using the method of Ref. 1. Overall the displacements are

[1] A. G. Every and K. Y. Kim, J. Acoust. Soc. Am. **95**, 2505 (1994).
[2] K. Y. Kim, A. G. Every, and W. Sachse, J. Acoust. Soc. Am. **95**, 1942 (1994).
[3] C. Y. Wang and J. D. Achenbach, Wave Motion **16**, 389 (1992); **18**, 273 (1993); Proc. R. Soc. London, Ser. A **449**, 441 (1995).
[4] C. Y. Wang and J. D. Achenbach, Wave Motion **24**, 227 (1996).
[5] A. Mourad and M. Deschamps, J. Acoust. Soc. Am. **97**, 3194 (1995); Acustica united with Acta Acustica **82**, 839 (1996).
[6] P. Borejko and S. K. Datta, Z. Angew. Math. Mech. **74**, T272 (1994).
[7] R. L. Weaver, W. Sachse, and K. Y. Kim, J. Appl. Mech. **63**, 337 (1996).
[8] H. Zhu, ASME J. Appl. Mech. **59**, S96 (1992).
[9] A. N. Norris, Proc. R. Soc. London, Ser. A **447**, 175 (1994).
[10] R. Burridge, P. Chadwick, and A. N. Norris, Proc. R. Soc. London, Ser. A **440**, 655 (1993).
[11] S. Tamura and M. Yagi, Phys. Rev. B **49**, 17378 (1994).
[12] V. K. Tewary and C. M. Fortunko, J. Acoust. Soc. Am. **91**, 1888 (1992).
[13] V. K. Tewary, Phys. Rev. B **51**, 15695 (1995).
[14] V. K. Tewary and C. M. Fortunko, J. Acoust. Soc. Am. **100**, 86 (1996).
[15] A. G. Every, W. Sachse, K. Y. Kim, and M. O. Thompson, Phys. Rev. Lett. **65**, 1446 (1990); A. G. Every and W. Sachse, Phys. Rev. B **44**, 6689 (1991).

[16] B. Castagnede and A. Mourad, C. R. Acad. Sci. Paris II **314**, 1301 (1992); II **316** (1993).

[17] J. F. Chai and T. T. Wu, J. Acoust. Soc. Am. **95**, 3232 (1994).

[18] Al. A. Kolomenskii and A. A. Maznev, Phys. Rev. B **48**, 14 502 (1993).

[19] R. L. Weaver, M. R. Hauser, and J. P. Wolfe, Z. Phys. B **90**, 27 (1993).

[20] R. E. Vines, S. Tamura, and J. P. Wolfe, Phys. Rev. Lett. **74**, 2729 (1995); R. E. Vines, M. R. Hauser, and J. P. Wolfe, Z. Phys. B **98**, 255 (1995).

[21] J. P. Wolfe and M. R. Hauser, Ann. Physik **4**, 99 (1995).

[22] J. Wesner, K. U. Würz, K. Hillmann, and W. Grill, in *Phonon Scattering in Condensed Matter VII*, edited by M. Meissner and R. O. Pohl (Springer-Verlag, Berlin, 1993), 68 pp.

[23] K. Y. Kim, W. Sachse, and A. G. Every, Phys. Rev. Lett. **70**, 3443 (1993).

[24] K. Y. Kim, A. G. Every, and W. Sachse, Int. J. Mod. Phys. B **8**, 2327 (1994).

[25] M. J. P. Musgrave and R. G. Payton, Q. J. Mech. Appl. Math. **34**, 235 (1982); **35**, 173 (1982).

[26] H. Lamb, Philos. Trans. R. Soc. London, Ser. A **203**, 1–42 (1904).

[27] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, Amsterdam, 1973).

[28] A. C. Eringen and E. S. Suhubi, *Elastodynamics* (Academic, New York, 1975).

[29] R. Burridge, Q. J. Mech. Appl. Math. **24**, 81 (1971).

[30] A. A. Maznev and A. G. Every, Int. J. Eng. Sci. **35**, 321 (1997).

[31] J. H. M. T. van der Hijden, *Propagation of Transient Elastic Waves in Stratified Anisotropic Media* (North-Holland, Amsterdam, 1987).

[32] J. R. Willis, Philos. Trans. R. Soc. London, Ser. A **274**, 435 (1973).

[33] R. G. Payton, *Elastic Wave Propagation in Transversely Isotropic Media* (Martinus Nijhoff, The Hague, 1983).

[34] G. W. Farnell, in *Physical Acoustics,* edited by W. P. Mason and R. N. Thurston (Academic, New York, 1970), Vol. VI, 109 pp.

[35] J. Lothe and D. M. Barnett, J. Appl. Phys. **47**, 428 (1976).

[36] H. Shirasaki and T. Makimoto, J. Appl. Phys. **49**, 658 (1978); **49**, 661 (1978); **50**, 2795 (1979).

[37] S. Tamura and K. Hongo, Jpn. J. Appl. Phys. **20**, Suppl. 20-3, 17 (1981).

[38] R. E. Camley and A. A. Maradudin, Phys. Rev. B **27**, 1959 (1983).

[39] L. A. Chernozatonskii and V. V. Novikov, Solid State Commun. **51**, 643 (1984).

[40] C. Höss and H. Kinder, Physica B **219&220**, 706 (1996).

[41] A. A. Maznev and A. G. Every, Acustica united with Acta Acustica **1**, 137 (1994); Solid State Commun. **97**, 679 (1996).

[42] K. Y. Kim, K. C. Bretz, A. G. Every, and W. Sachse, J. Appl. Phys. **79**, 1857 (1996).

[43] A. A. Maznev, A. A. Kolomenskii, and P. Hess, Phys. Rev. Lett. **75**, 332 (1995).

[44] A. G. Every, G. L. Koos, and J. P. Wolfe, Phys. Rev. B **29**, 2190 (1984).

[45] D. C. Hurley, A. G. Every, and J. P. Wolfe, J. Phys. C **17**, 3157 (1984).

[46] A. G. Every, Phys. Rev. B **33**, 2719 (1986).

[47] G. Carlotti, D. Fioretto, L. Giovanni, F. Nizzoli, G. Socino, and L. Verdini, J. Phys.: Condens. Matter **4**, 257 (1992).

[48] R. E. Camley and F. Nizzoli, J. Phys. C: Solid State Phys. **18**, 4795 (1985).

[49] K. Y. Kim and W. Sachse, J. Appl. Phys. **75**, 1435 (1994).

[50] K. Y. Kim, Wave Motion **20**, 83 (1994).

Blocking property of energy vortices in elastic waveguides

Evgeny Glushkov[a)] and Natalya Glushkova

Kuban State University, P.O. Box 4102, Krasnodar 350080, Russia

Energy vortices in time-averaged energy flows of time-harmonic fields are considered. The purpose of the paper is to verify the supposition that the energy flux along an elastic waveguide with an obstacle is blocked completely at the stop frequencies by the vortices. It is hoped that this work will draw attention to the analogy between energy and fluid flows which is potentially fruitful for understanding various wave phenomena. The normal-mode diffraction by a surface punch on an elastic layer is taken as an example. For tracing energy streamlines in the near field up to the obstacle, a special efficient semianalytical approach has been developed. To illustrate the stated supposition, plots of transmission coefficients versus frequency and figures of the near-field streamline structures are given. © *1997 Acoustical Society of America.* [S0001-4966(97)06108-0]

PACS numbers: 43.20.Gp, 43.20.Mv, 43.20.Wd [JEG]

INTRODUCTION

A time-averaged energy flux E across a surface S in the time-harmonic elastic wave field $\mathbf{u}(\mathbf{x})e^{-i\omega t}$ is introduced as an integral of the normal to S component e_n of the averaged power density vector (Umov–Poynting vector) $\mathbf{e}(\mathbf{x})$ over S:

$$E = \int\int_S e_n(\mathbf{x})dS.$$

Here, $e_n = (\mathbf{e},\mathbf{n}) = -(\omega/2)\text{Im}(\mathbf{u},\tau_n)$, where $\mathbf{n}$ is the unit outward normal for S and τ_n is the corresponding stress vector. Continuous vector field $\mathbf{e}(\mathbf{x}) = \{e_x, e_y, e_z\}$ determines in the space $\mathbf{x} = \{x,y,z\}$ a set of noncrossing curves $\mathbf{x}(s)$ to which $\mathbf{e}(\mathbf{x})$ is tangential at any point:

$$\frac{d\mathbf{x}}{ds} = \frac{\mathbf{e}(\mathbf{x})}{|\mathbf{e}(\mathbf{x})|}. \tag{1}$$

Every curve passing through a given point $\mathbf{x}_0$,

$$\mathbf{x}(0) = \mathbf{x}_0, \tag{2}$$

is determined uniquely from Cauchy problems (1) and (2). These curves are known as the energy streamlines,[1] although such a name is not quite correct for the instantaneous energy transfer is not described by the averaged vector $\mathbf{e}$ and, consequently, real power flow differs from those lines at any fixed time.[2] Nevertheless, the time-averaged flow provides more obvious and important visual information about total energy transfer in waveguides than a large number of the instantaneous values. Similar to the velocity field in liquid, the time-averaged energy flow is subjected to the conservation law ($E \equiv 0$ for a closed S without sources and run-off inside) and to the equation of continuity div $\mathbf{e} = 0$ (it follows from the equilibrium equations in stress components). Therefore the analogy with a fluid flow, also described by the streamlines, seems to be appropriate, together with stating the problem of developing the ''fluid dynamics'' of energy fluxes in solids. If general laws and properties of the energy–liquid are formulated, it would give new insight into the nature of various wave phenomena.

It has been known[1] that there exist zones in the time-averaged energy flows in which the streamlines are closed. Centers of these zones, named energy vortices, are points of wavefront dislocations[3] at which $\mathbf{e} \equiv 0$.[4] We have been studying the energy flows and vortices in isotropic elastic layered waveguides,[5–9] trying, first of all, to clear up the role of the vortices in various resonance phenomena. These works obtained numerical results demonstrating a possibility of backward energy fluxes coming from infinity (energy circulation through infinity) with considerable energy accumulation in such vortices near the resonance frequencies.[6,7] Those vortices, considered mainly in the far field, were the result of normal-mode interference. More interesting, but simultaneously more difficult mathematically, was the study of vortices arising in the near field of a source or an obstacle as a result of diffraction. To trace the streamlines in the near field close to the object required the use of a strict and numerically accurate solution of a mixed boundary-value problem instead of the far-field asymptotics. Therefore, a special semianalytical approach for solving the mixed problems related to diffraction by surface and inner horizontal and vertical obstacles in layered elastic waveguides has been developed.[9–11]

When analyzing the near-field streamline structure under a rigid punch vibrating on an elastic layer, it was shown that vortices, changing size with the remaining various free space for outflow of energy emitted into a waveguide, in fact, controlled the important source characteristic, energy efficiency.[8,9] Specifically, the nature of the so-called high-frequency punch-layer resonances[12] had become clearer when it was demonstrated that they occurred if the vortices completely blocked the energy outflow from under the punch.

This property of the vortices, to dam the energy flow partly or completely, seems to be the most interesting of its manifestations. For example, the directions between the lobes in radiating and scattering diagrams are probably none other than the shadow of the vortices arising near a radiator[13,14] or a scatterer. In the waveguides such near-field

[a)]Electronic mail: evg@kgu.kuban.su

vortices, resulting from normal-mode diffraction by surface steps, punches, cracks, inclusions, joint lines of semistrips, and so on, become apparent in regulating the transmission and reflection coefficients even to the point of completely locking up the waveguide at a stop frequency, when the wave power does not pass by an obstacle through the, one can say, free cross section. The purpose of the present paper is to give an example of such power blocking at the stop frequency, to draw the attention of specialists of fluid dynamics intuition to the energy flow properties.

Earlier, when numerically analyzing vibration of a group of punches on an elastic layered substrate subjected to the surface waves, there was noted a sharp amplitude decrease of the second punch at the resonance frequency of the front one, supposedly, because the vortices near the latter grew maximally in that case. Therefore the rigid punch on an elastic layer model was chosen to analyze the streamline structure at the stop frequency in detail.

To obtain proper numerical results in the near field one has to use not only traveling modes but also modes related to the complex roots of the dispersion equation up to very large (hundreds of thousands) mode numbers. In addition to the large scale, direct truncation of the infinite linear algebraic systems arising here is numerically unstable. Regularization of the systems is ensured by taking into account the asymptotic behavior of unknown constants t_k, as $k \to \infty$, that depend on the singularity at the edges. This approach developed earlier for the axisymmetric contact problems[10] is modified for the plane case considered here. It is worth noting that it is harder to carry out numerical analysis in the plane case due to the slower rate of series convergence, especially for the stress components under the edges.

As for vertical obstacles (surface steps, groves, compound waveguides, etc.) the main idea of regularization[9] is based on the generalized orthogonality of the normal modes.[15] Numerical examples of the energy transfer over steps and across joint lines are given in Ref. 11.

I. MATHEMATICAL STATEMENT AND TECHNIQUE

Let consider an elastic layer of thickness h ($-\infty \leqslant x \leqslant \infty, -h \leqslant z \leqslant 0$) fixed from below ($\mathbf{u}=0, z=-h$) with a rigid strip punch of width $2a$ contacting frictionlessly with the stress-free upper surface $z=0$:

$$\sigma_z = \begin{cases} q(x), & |x| < a, \\ 0, & |x| > a, \end{cases}$$

$$\tau_{xz} = 0, \quad -\infty \leqslant x \leqslant \infty, \qquad (3)$$

$$u_z = w(x), \quad |x| < a.$$

Here $q(x)$ and $w(x)$ are the unknown contact stress and vertical displacement of the punch, $\mathbf{u} = \{u_x, u_z\}$; the harmonic factor $e^{-i\omega t}$ is omitted here and thereafter. For simplicity sake it is quite enough to restrict our consideration to the special case of unmoved (fixed) punch: $w \equiv 0$.

The punch is subjected to an incident wave field $\mathbf{u}_0(x,z)$ coming from the left, from infinity, or from a not distant source. Wave field $\mathbf{u}$ in the layer is a combination of the given field $\mathbf{u}_0$ and scattered field $\mathbf{u}_1$: $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_1$. The

latter is expressed by way of an integral in terms of the Green's matrix components and unknown contact stress $q(x)$:[6,7]

$$\mathbf{u}_1(x,z) = \frac{1}{2\pi} \int_\Gamma \mathbf{b}(\alpha,z) Q(\alpha) e^{-i\alpha x} \, d\alpha, \qquad (4)$$

where

$$Q(\alpha) = \int_{-a}^{a} q(x) e^{i\alpha x} \, dx, \qquad (5)$$

$\mathbf{b} = \{-i\alpha P_1, R_1\}^T/\Delta$ is the second column of the two-dimensional (2-D) Green's matrix Fourier symbol; for the considered layer stuck to the rigid base, the form of functions $P_1(\alpha,z)$, $R_1(\alpha,z)$, $\Delta(\alpha)$ is given in Ref. 10.

Integration path Γ goes in the complex plane α along the real axis, deviating from it in the upper or lower half plane in accordance with the principle of limiting absorption to bypass real singular points ζ_k, $k=1,2,...,N_R$. Hereafter, z_l, ζ_k are zeros and poles of meromorphic function

$$K(\alpha) = R_1(\alpha,0)/\Delta(\alpha)$$

[roots of $R_1(\alpha,0)=0$ and $\Delta(\alpha)=0$, respectively] located in the α plane above Γ. A finite number of the first $\pm z_l$, $\pm \zeta_k$ lies on the real axis; with large numbers l,k their location is described by the asymptotics:

$$z_l \sim i\pi(m+\tfrac{1}{4}) \pm \tfrac{1}{2}\ln(4\pi(m+\tfrac{1}{4})/(3-4\nu)), \quad m=[l/2] \to \infty, \qquad (6)$$

$$\zeta_k \sim i\pi n \pm \ln(2\pi n/\sqrt{3-4\nu}), \quad n=[k/2] \to \infty,$$

where ν is the Poisson ratio.

Closing Γ in the low half-plane α for $x \geqslant a$ and in the upper one for $x \leqslant -a$ and using the residual Cauchy theorem transform Eq. (4) into

$$\mathbf{u}_1^\pm(x,z) = \sum_{k=1}^{\infty} t_k^\pm \mathbf{a}_k^\pm(z) e^{i\zeta_k(\pm x - a)}, \quad |x| > a, \qquad (7)$$

where

$$t_k^\pm = i r_k Q(\mp \zeta_k) e^{i\zeta_k a},$$

$$r_k = \operatorname{Res} K(\alpha)|_{\alpha=\zeta_k} = R_1(\zeta_k,0)/\Delta'(\zeta_k), \qquad (8)$$

$$\mathbf{a}_k^\pm(z) = \mathbf{a}(\mp \zeta_k, z), \quad \mathbf{a}(\alpha,z) = \mathbf{b}(\alpha,z)/K(\alpha).$$

Here the upper and lower superscripts correspond to $x > a$ and $x < -a$, respectively, i.e., $\mathbf{u}_1^+$ and $\mathbf{u}_1^-$ are the transmitted and reflected parts of the scattered field $\mathbf{u}_1$.

As for the field under the punch, zeros z_l become noneliminable poles of $Q^\pm(\alpha)$ due to splitting $Q(\alpha) = Q^+ e^{i\alpha a} + Q^- e^{-i\alpha a}$ required for the path closing, so that in the case of a fixed punch only zeros $\{z_l\}_{l=1}^{\infty}$ contribute in the final representation:

E. Glushkov and N. Glushkova: Blocking in elastic waveguides

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_0 = \sum_{l=1}^{\infty} \; [c_l^+ \mathbf{b}_l^+(z)e^{iz_l(a-x)}$$

$$+ c_l^- \mathbf{b}_l^-(z)e^{iz_l(a+x)}], \quad |x| < a,$$

$$\mathbf{b}_l^{\pm}(z) = \mathbf{b}(\pm z_l, z), \quad c_l^{\pm} \text{ see below (17).} \tag{9}$$

Unknown $q(x)$ satisfies the Wiener–Hopf type integral equation following from Eq. (4) and boundary condition (3):

$$\mathcal{K}q \equiv \int_{-a}^{a} k(x-\xi)q(\xi)d\xi = w(x) - u_{0,z}(x) \overset{\text{Def}}{\equiv} f(x), \quad |x| < a \tag{10}$$

with kernel $k(x) = 1/2\pi \int_{\Gamma} K(\alpha)e^{-i\alpha x}\, d\alpha$.

Equation (10) is brought by the Fourier transform

$$\mathcal{F}[u] = \int_{-\infty}^{\infty} u(x)e^{i\alpha x}\, dx \equiv U(\alpha)$$

to the Wiener–Hopf functional equality

$$K(\alpha)Q(\alpha) = F(\alpha) + \Phi(\alpha), \tag{11}$$

where $F(\alpha) = \mathcal{F}[f]$ and

$$\Phi(\alpha) = \mathcal{F}[u_{1,z}(x,0), |x| > a]$$

$$= i \sum_{k=1}^{\infty} \left[\frac{t_k^+ e^{ia\alpha}}{\alpha + \zeta_k} - \frac{t_k^- e^{-ia\alpha}}{\alpha - \zeta_k} \right]. \tag{12}$$

The requirement that poles $\pm z_l$ of the entire function $Q(\alpha)$ induced by equality

$$Q = (F + \Phi)/K \tag{13}$$

must be eliminable $[F(\pm z_l) + \Phi(\pm z_l) = 0]$ yields an infinite algebraic system relative to $t_k^{\pm}$, which in turn is split into two independent systems by substituting $t_k^{\pm} = (s_k^+ \pm s_k^-)/2$:

$$\sum_{k=1}^{\infty} d_{lk}^{\pm} s_k^{\pm} = f_l^{\pm}, \quad l = 1,2,3,\dots,$$

$$d_{lk}^{\pm} = i\left(\mp \frac{1}{z_l - \zeta_k} + \frac{e^{2iaz_l}}{z_l + \zeta_k} \right), \tag{14}$$

$$f_l^{\pm} = -(F(z_l) \pm F(-z_l))e^{iaz_l}.$$

Direct truncation of systems (14) is numerically unstable. But replacing constants $s_k^{\pm}$ by their asymptotics as $k \to \infty$ starting from a certain number $k = N+1$ results in the regularized asymptotically equivalent stable finite systems. The required asymptotics of $s_k^{\pm}$, in view of Eq. (5), can be expressed as the asymptotics of $Q(\pm \zeta_k)$ as $|\zeta_k| \to \infty$, i.e., the asymptotics of oscillate integrals. Specifically, root behavior at the edges

$$q(x) \sim \sum_{m=1}^{M} q_m^{\pm} \cdot (a \mp x)^{m-3/2}, \quad x \to \pm a, \tag{15}$$

contributes as

$$Q(\mp \zeta_k) \sim e^{-ia\zeta_k} \sum_{m=1}^{M} q_m^{\pm}/(-i\zeta_k)^{m-1/2},$$

$$|\zeta_k| \to \infty, \quad \text{Im } \zeta_k > 0.$$

Earlier, in the axisymmetrical problems, it was quite enough to use only the first term of the expansion to stabilize the numerical procedure. However, in the plane case the stabilization required is not achieved even with M terms of the expansion (15) for it leaves aside the oscillation terms of $q(x)$ contributing as $e^{-i\zeta_k a}O(|\zeta_k|^{-1})$ as $|\zeta_k| \to \infty$.

Indeed, based upon Eq. (13), $q(x)$ can be represented in the same way as $\mathbf{u}_1$ in terms of residuals at the induced poles z_l:

$$q(x) = \sum_{n=1}^{\infty} \; [c_n^+ e^{iz_n(a-x)} + c_n^- e^{iz_n(a+x)}], \tag{16}$$

$$c_n^{\pm} = p_n\left[iF^{\pm}(\pm z_n) - \sum_{k=1}^{\infty} \frac{t_k^{\pm}}{z_n + \zeta_k} \right],$$

$$p_n = \text{Res } K^{-1}(\alpha)|_{\alpha = z_n}, \tag{17}$$

$$F^{\pm}(\alpha): \quad F(\alpha) = F^+(\alpha)e^{ia\alpha} + F^-(\alpha)e^{-ia\alpha}.$$

Consequently,

$$Q(\alpha) = 2\sum_{n=1}^{\infty} \left[c_n^+ \frac{\sin(\alpha - z_n)a}{\alpha - z_n} + c_n^- \frac{\sin(\alpha + z_n)a}{\alpha + z_n} \right]e^{iz_n a}. \tag{18}$$

The first several terms in Eq. (16) corresponding to real z_n result in oscillation of $q(x)$, the remaining terms, due to (6), diminish exponentially from the edges, but the series diverges at $x = \pm a$ yielding the root singularity of the form (15). That is why the optimal way to stabilize the systems is to keep a finite number of the terms in (16) and to replace the rest by the root expansion (15). This leads to asymptotically equivalent stable systems

$$\sum_{k=1}^{N} d_{lk}^{\pm} s_k^{\pm} + \sum_{n=1}^{M_1} a_{ln}^{\pm}(c_n^+ \pm c_n^-) + \sum_{m=1}^{M_2} b_{lm}^{\pm}(q_m^+ \pm q_m^-) = f_l^{\pm},$$

$$l = 1,2,\dots,N + M_1 + M_2, \tag{19}$$

$$a_{ln}^{\pm} = - \sum_{k=N+1}^{\infty} d_{lk}^{\pm} r_k\left(\frac{1}{z_n + \zeta_k} \mp \frac{e^{2iz_n a}}{z_n - \zeta_k} \right),$$

$$b_{lm}^{\pm} = i \sum_{k=N+1}^{\infty} \frac{d_{lk}^{\pm} r_k}{(-i\zeta_k)^{m-1/2}}$$

relative to $N + M_1 + M_2$ unknowns $s_k^{\pm}$, $c_n^{\pm}$, $q_m^{\pm}$.

It should be noted that there exists another way of regularization: substitution of (18) in equality (8) taking into account (17) leads to the Fredholm system of the second kind. That system in its structure and properties is like the one obtained in the axisymmeric case by extracting and inverting explicitly the main part of the matrix operator: $\|1/(z_l - \zeta_k)\|_{l,k=1}^{\infty}$.[16] However, such an approach turns out to be more cumbersome on account of the double series and, in addition, it remains unstable near the frequencies at which $\pm z_l$, passing from the real axis into the complex plane, merged in double zeros.

The described scheme is easily generalized for a group of horizontal obstacles (e.g., strip electrodes on the elastic substrate). Generalization for the case of matrix integral op-
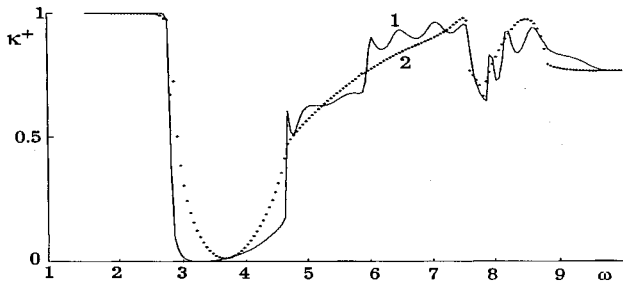
FIG. 1. Transmission coefficient κ^+ vs ω for the first mode; $\nu=0.3$, a/h $=2$ (solid curve) and 0.2 (dash).

FIG. 3. Power distribution coefficients $\kappa_0^\pm$ versus ω ($a=2$, $x_0=-5$).

erators (stuck punches, horizontal cracks) is carried out in the same way as in the axisymmetrical problem.[10]

II. STREAMLINE STRUCTURE NEAR THE OBSTACLE

The stop frequency phenomenon implies that the transmission coefficient of a waveguide becomes zero. The transmission and reflection coefficients κ^+, κ^- are ratios $\kappa^\pm$ $=E^\pm/E_0$, where E_0 is the time-averaged power transferred by the incident field $\mathbf{u}_0$ through the layer cross section, E^+ is the part of the power bypassing the obstacle and flowing to $+\infty$, and E^- is the power transported to $-\infty$ by the reflected waves. Values E^+, E^- are obtained by integrating the power density component e_x expressed via $\mathbf{u}=\mathbf{u}_0+\mathbf{u}_1$ over the layer cross section $-h\leq z\leq 0$ with $\mathbf{n}=\mathbf{u}_x=\{1,0\}$, $x>a$ and $\mathbf{n}=-\mathbf{n}_x$, $x<-a$, respectively. Since $E^\pm$, in fact, are independent of x, employing the far-field asymptotics at $|x|\gg 1$, expressed in terms of traveling modes only, is the most convenient here. The latter relates to the real poles ζ_k, k $=1,2,...,N_R$. For the layer considered, the real branches of dispersion curves $\zeta_k(\omega)$ issue from the points $\omega_k=1$, 57, 2.94, 4.71, 7.85, 8.82,... of axis ω (dimensionless $\omega=2\pi f h/c_S$, c_S is a shear wave speed, f is a frequency here).

If $\mathbf{u}_0$ comes from $-\infty$, it consists of traveling terms only. The energy conservation law implies the power balance condition $E_0=E^++E^-$ and, consequently, the equality κ^+ $+\kappa^-=1$. Satisfaction of this condition served as an additional control on the numerical results (the main and most difficult control was the numerical check of the field continuity under the edges $x=\pm a$, $-h\leq z\leq 0$).

Two examples of the transmission coefficient versus frequency dependence for the first mode $\mathbf{u}_0=\mathbf{a}_1^+(z)e^{i\zeta_1 x}$ and the fixed punch ($w=0$) of width $a/h=2$ and 0.2, $\nu=0.3$ are

given in Fig. 1. One can see in Fig. 1 the stop frequencies at $\omega\approx 3.3$ and $\omega\approx 3.5$ whose locations are almost independent of the size of the obstacle. Apparently, power transfer is governed here primarily by the diffraction by the front edge $x=-a$. From the energy–liquid viewpoint, the diffraction gives rise to energy vortices whose size and location control power transmission through the cross section adjacent to the obstacle. At the stop frequency we have received closed streamlines only, i.e., power flow in the layer has fallen completely into a countable set of vortices. Integration of e_x over any cross section yielded zero power flux in that case. Therefore field $\mathbf{u}_0$ coming from a not distant source seemed to be

FIG. 2. Wave fields' geometry with a not distant source.
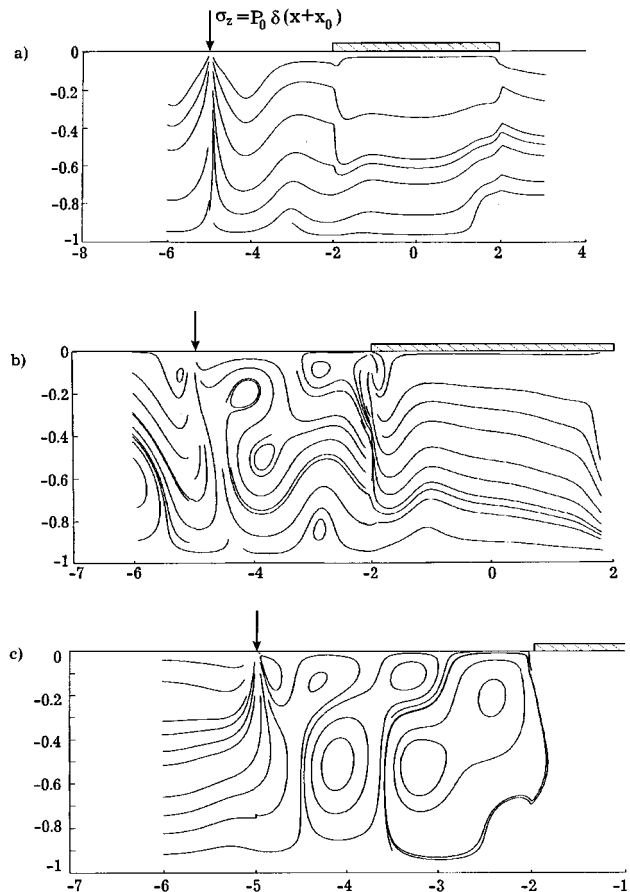
FIG. 4. Near-field time-averaged energy streamline structure (a) without vortices in the case of almost full transmission ($\omega=2.8$), (b) in the case of partial reflection ($\omega=3.0$), and (c) at the stop frequency ($\omega=3.6$).

more physically appropriate for examining streamline structure at the stop frequency.

Consider point surface loading $\sigma_z = p_0 \delta(x - x_0)$, $x_0 < -a$ as an example (Fig. 2). Total power

$$E_0 = \frac{\omega}{2} \left\{ |p_0|^2 \sum_{k=1}^{N_R} r_k + \mathrm{Im}\left[p_0^* \sum_{k=1}^{\infty} t_k^- e^{-i\zeta_k(x_0+a)} \right] \right\}$$

emitted by the source into the layer is divided into two parts transported at the left of the source to $-\infty$ and at the right to the obstacle. In this case $\kappa_0^\pm = E^\pm/E_0$ should be referred to as power distribution coefficients. Naturally, the energy balance condition also leads to equality $\kappa_0^+ + \kappa_0^- = 1$ here. Note that the reflected field $\mathbf{u}_1^-$ influences E_0 as a function of distance $-(a + x_0)$ and frequency ω, so that sometimes $E_0 = 0$ (the source is locked). But our interest is the case of $\kappa_0^+ = 0$ with $E_0 \neq 0$.

Figure 3 displaying $\kappa_0^\pm(\omega)$ (with $x_0 = -5$, $a = 2$) shows approximately the same stop frequency as with the mode coming from infinity above. Figure 4 illustrates how the vortices, growing near the front edge, block the energy flux, to the point of total reflection and turning of the emitted energy flow entirely to the left. The stop frequency $\omega = 3.6$ occurs here after the first inverse mode range $2.88 < \omega < 2.94$ with two traveling modes excited. In the one mode range $1.57 < \omega < 2.88$ there are no vortices near the obstacle, therefore, flux passes by unobstructed and $\kappa_0^+ \approx \kappa_0^-$ [Fig. 4(a), $\omega = 2.8$]. The vortices appear for $\omega > 2.88$ when the next traveling modes are added; a typical streamline picture is shown in Fig. 4(b) ($\omega = 3$). Notice that the vortices resulting from the interference are found to the left of the source also, so that sometimes it occurs $\kappa_0^+ > \kappa_0^-$ (Fig. 3). At the stop frequency $\omega = 3.6$ all energy emitted by the source turns to the left, being dammed up from the right side by the vortices [Fig. 4(c)].

ACKNOWLEDGMENTS

[1] R. V. Waterhouse, T. W. Yates, D. Feit, and Y. N. Liu, ''Energy streamlines of a sound source,'' J. Acoust. Soc. Am. **78**, 758–762 (1985).

[2] J. A. Mann III, J. Tichy, and A. J. Romano, ''Instantaneous and time-averaged energy transfer in acoustic fields,'' J. Acoust. Soc. Am. **82**, 17–29 (1987).

[3] V. A. Zhuravlev, I. K. Kobozev, and Yu. A. Kravtsov, ''Energy fluxes near dislocation of the phase field of the wave front,'' Zh. Eksp. Teor. Fiz. **104**, 3769–3783 (1993) (in Russian).

[4] R. V. Waterhouse, D. G. Crighton, and J. E. Ffowcs-Williams, ''A criterion for an energy vortex in a sound field,'' J. Acoust. Soc. Am. **81**, 1323–1326 (1987).

[5] V. A. Babeshko, E. V. Glushkov, and N. V. Glushkova, ''Analysis of wave fields generated in a stratified elastic half-space by surface sources,'' Sov. Phys. Acoust. **32**, 223–226 (1986).

[6] V. A. Babeshko, E. V. Glushkov, and J. F. Zinchenko, *Dynamics of the Inhomogeneous Linear-elastic Media* (Nauka, Moscow, 1989), p. 344 (in Russian).

[7] V. A. Babeshko, E. V. Glushkov, and N. V. Glushkova, ''Energy vortices and backward fluxes in elastic waveguides,'' Wave Motion **16**, 183–192 (1992).

[8] V. A. Babeshko, E. V. Glushkov, N. V. Glushkova, and E. V. Kirillova, ''Energy localisation under high-frequency resonance conditions,'' Sov. Phys. Dokl. **37**, 443–444 (1992).

[9] E. V. Glushkov, N. V. Glushkova, and E. V. Kirillova, ''Normal mode diffraction in elastic layered waveguides; resonances and energy vortices,'' in *Proceedings of the Third International Conference on Mathematical and Numerical Aspects of Wave Propagation, Mandelieu-La-Napoule, France*, 1995, edited by Gary Cohen (SIAM, Philadelphia, 1995), pp. 604–612.

[10] E. V. Glushkov, N. V. Glushkova, and E. V. Kirillova, ''The dynamic contact problem for a circular punch adhering to an elastic layer,'' J. Appl. Math. Mech. **56**, 675–679 (1992).

[11] E. V. Glushkov, N. V. Glushkova, and O. N. Lapina, ''Normal mode diffraction in compound and step elastic waveguides,'' submitted to Prikl. Math. Mekh. (in Russian).

[12] V. A. Babeshko, I. I. Vorovich, and I. F. Obraztsov, ''A high-frequency resonance phenomenon in semirestricted media with inhomogeneuties,'' Izv. Akad. Nauk SSSR, Mekh. Tverdogo Tela (Proceedings of the USSR Academi of Science. Solid Mechanics) **3**, 74–83 (1990) (in Russian).

[13] G. Krishnappa and J. M. McDougall, ''Sound intensity distribution and energy flow in the nearfield of a clamped circular plate,'' J. Vib. Acoust., Stress, Reliability Design **111**, 465–471 (1989).

[14] R. J. Weizman, and E. V. Zinoviev, ''On sound energy flow caused by plate vibration,'' J. Sound Vib. **173**, 429–432 (1994).

[15] A. S. Zilbergleit and B. M. Nuller, ''Generalised orthogonality of homogeneous solutions in elastodynamics,'' Dokl. Akad. Nauk SSSR **234**, 333–335 (1977) (in Russian).

[16] E. V. Glushkov and N. V. Glushkova, ''Determination of the dynamic contact stiffness of an elastic layer,'' Prikl. Math. Mekh. U.S.S.R. **54**, 393–397 (1990).

Doubly asymptotic approximations for transient poroelastodynamics[a)]

Quan Qi[b)] and Thomas L. Geers

Center for Acoustics, Mechanics and Materials, Department of Mechanical Engineering,
University of Colorado, Boulder, Colorado 80309-0427

A doubly asymptotic approximation (DAA) is an approximate temporal impedance relation at the boundary of a continuous medium. Here, first- and second-order DAAs are formulated for an infinite external poroelastic medium described by Biot's equations. As with their acoustic, elastodynamic, and electromagnetic predecessors, the poroelastodynamic DAAs approach exactness in both the early-time (high-frequency) and late-time (low-frequency) limits, effecting a smooth transition between. They also lend themselves to straightforward boundary-element discretization, producing matrix ordinary differential equations in time that are readily solved by numerical integration. An initial examination of poroelastodynamic DAA accuracy is presented for two problems with spherical symmetry. © *1997 Acoustical Society of America.* [S0001-4966(97)04308-7]

PACS numbers: 43.20.Gp, 43.20.Tb, 43.20.Px [JEG]

INTRODUCTION

Treating the transient dynamic interaction between a body in contact with an infinite poroelastic medium is a challenging task. A doubly asymptotic approximation greatly simplifies the problem by providing a time-domain impedance relation for the medium at a desired impedance surface. This relation is temporally local but spatially nonlocal. Discretization by the boundary-element method yields ordinary differential equations with full, constant matrices that pertain only to degrees of freedom on the impedance surface. The impedance surface may coincide with the contact surface between the body and the medium, or it may be positioned away from the contact surface, thereby facilitating the treatment of local nonlinear behavior of the medium.

DAAs were first formulated in the 1970s to treat the external fluid–structure interaction in underwater shock problems (Geers, 1971, 1978). During the following decade, the acoustic DAA methodology was improved (Felippa, 1980; Geers and Felippa, 1983; Nicolas-Vullierme, 1991) and DAAs were formulated for elastodynamic and electromagnetic problems (Underwood and Geers, 1981; Mathews and Geers, 1987; Geers and Zhang, 1988). Recent enhancements include DAAs for internal acoustic domains (Geers and Zhang, 1994) and higher-order DAAs for external elastic domains (Wolf and Song, 1995; Geers and Lewis, 1997).

Biot's classical papers (Biot, 1956a, b) describe a theory for the dynamic behavior of a poroelastic medium in terms of interpenetrating solid and fluid phases. The theory has proven useful for treating a wide range of phenomena, such as fluid diffusion through an elastic matrix and wave dispersion/attenuation due to relative motion between fluid and solid phases (Johnson, 1984). The theory admits three wave types: a slow longitudinal wave, a fast longitudinal wave, and a shear wave. An interesting feature of the slow wave is that it is diffusive at low frequencies and propagatory at high frequencies. The wave behavior predicted by the theory has been confirmed by experiment (Johnson, 1984).

This paper presents the first formulation of DAAs for a poroelastic medium. Four surface-displacement variables are involved (Bonnet, 1987): three orthogonal displacements of the solid phase and the normal displacement of the fluid phase. The four conjugate surface-force variables are three orthogonal surface tractions on the solid phase and the product of porosity and pore pressure in the fluid phase. Approximate early-time (high-frequency) and late-time (low-frequency) relations between the conjugate sets are obtained through the use of generalized Helmholtz potentials (Norris, 1985; Zimmerman and Stern, 1993). First-order and second-order DAAs are then obtained by the method of operator matching (Nicolas-Vullierme, 1991; Geers and Zhang, 1994; Geers and Lewis, 1997). The extension to boundary-element forms is straightforward. Finally, an exploratory examination of DAA accuracy is carried out by comparing ''exact'' and DAA response solutions for a thin spherical shell embedded in a poroelastic medium and subjected to a step internal pressure. A second examination follows in which the shell is removed and the step pressure is applied directly to the medium.

I. POROELASTODYNAMIC EQUATIONS

Biot's original equations of motion for a poroelastic medium involve six response variables: three orthogonal displacements each for the solid and fluid phases. In 1985 Norris expressed those variables in a generalized Helmholtz decomposition of two scalar potentials and a vector potential. In 1987 Bonnet showed that four response variables are sufficient (three orthogonal displacements of the solid phase and the product of porosity and pore pressure in the fluid phase). Norris (1985) and Zimmerman and Stern (1993) determined the characteristics of the associated poroelastic waves, which form the basis of the DAA formulation.

A. Equations of motion

Biot's dynamical equations for a poroelastic medium are (Biot, 1956a, b)

$$\rho_{11}\frac{\partial^2}{\partial t^2}\mathbf{u}+\rho_{12}\frac{\partial^2}{\partial t^2}\mathbf{U}=\nabla\cdot\boldsymbol{\sigma}+b\frac{\partial}{\partial t}(\mathbf{U}-\mathbf{u})+\mathbf{B}^s,$$

$$\rho_{12}\frac{\partial^2}{\partial t^2}\mathbf{u}+\rho_{22}\frac{\partial^2}{\partial t^2}\mathbf{U}=\nabla(-q)-b\frac{\partial}{\partial t}(\mathbf{U}-\mathbf{u})+\mathbf{B}^f, \quad (1)$$

where $\mathbf{u}(\mathbf{x},t)$ and $\mathbf{U}(\mathbf{x},t)$ are the solid and fluid displacement vectors, respectively; $\boldsymbol{\sigma}(\mathbf{x},t)$ is the stress tensor in the solid phase; $q(\mathbf{x},t)=fp(\mathbf{x},t)$ is the product of medium porosity f and pore pressure in the fluid phase $p(\mathbf{x},t)$; $\mathbf{B}^s(\mathbf{x},t)$ and $\mathbf{B}^f(\mathbf{x},t)$ are body forces per unit volume applied to the solid and fluid phases, respectively; b is the viscous coupling parameter that relates drag force to relative velocity between the solid and fluid phases; ρ_{11}, ρ_{12}, and ρ_{22} are Biot mass densities (Appendix A). His stress-displacement relations are

$$\boldsymbol{\sigma}=(P-2N)(\nabla\cdot\mathbf{u})\mathbf{I}+N(\nabla\mathbf{u}+\nabla\mathbf{u}^T)+Q(\nabla\cdot\mathbf{U})\mathbf{I},$$

$$-q=Q\nabla\cdot\mathbf{u}+R\nabla\cdot\mathbf{U}, \quad (2)$$

where N is the shear modulus of the solid phase, P, Q, and R are Biot moduli (Appendix A), and $\mathbf{I}$ is the second-order identity tensor. The total stress in the poroelastic medium is given by $\boldsymbol{\sigma}^t=\boldsymbol{\sigma}-q\mathbf{I}$.

It is convenient to reduce the six displacement variables in (1) to four response variables (Bonnet, 1987). To achieve this, we Laplace transform (1) and (2) and eliminate solid stress and fluid displacement. This yields the following three equations of motion:

$$(\lambda+N)\nabla(\nabla\cdot\widetilde{\mathbf{u}})+N\nabla^2\widetilde{\mathbf{u}}-\bar\rho s^2\widetilde{\mathbf{u}}+\alpha\nabla(\widetilde{q})=-\widetilde{\mathbf{F}}, \quad (3)$$

in which a tilde indicates a Laplace-transformed variable and

$$\lambda=P-2N-Q^2/R, \quad \bar\rho=(\bar\rho_{11}-\bar\rho_{12}^2/\bar\rho_{22}),$$

$$\alpha=(\bar\rho_{12}/\bar\rho_{22})-Q/R, \quad \widetilde{\mathbf{F}}=\widetilde{\mathbf{B}}^s-(\bar\rho_{12}/\bar\rho_{22})\widetilde{\mathbf{B}}^f, \quad (4)$$

where the modified (s dependent) densities are given by $\bar\rho_{ij}=\rho_{ij}+(-1)^{i+j}(b/s)$, $i,j=1,2$. The fourth equation of motion is obtained by eliminating fluid displacement from the second equation in Laplace-transformed (1) and (2); this yields

$$\nabla^2\widetilde{q}-(\bar\rho_{22}/R)s^2\widetilde{q}+\bar\rho_{22}s^2\alpha\nabla\cdot\widetilde{\mathbf{u}}=-\widetilde{V}, \quad (5)$$

where $\widetilde{V}=-\nabla\cdot\widetilde{\mathbf{B}}^f$.

The preceding formulation differs slightly from that of Zimmerman and Stern (1993), in which response variables defined in the transform domain are not simply the transforms of the corresponding variables in the time domain.

B. Poroelastodynamic potentials

In the absence of body forces, $\mathbf{u}(\mathbf{x},t)$ and $q(\mathbf{x},t)$ may be expressed as generalized Helmholtz decompositions of two scalar potentials and a vector potential (Appendix A):

$$\widetilde{\mathbf{u}}(\mathbf{x},s)=\nabla\widetilde{\phi}_f(\mathbf{x},s)+\nabla\widetilde{\phi}_s(\mathbf{x},s)+\nabla\times\nabla\times\widetilde{\boldsymbol{\psi}}(\mathbf{x},s),$$

$$\widetilde{q}=A_f\nabla^2\widetilde{\phi}_f(\mathbf{x},s)+A_s\nabla^2\widetilde{\phi}_s(\mathbf{x},s), \quad (6)$$

where $\widetilde{\phi}_f$ and $\widetilde{\phi}_s$ are Laplace-transformed scalar potentials, respectively, and $\widetilde{\boldsymbol{\psi}}$ is a Laplace-transformed vector potential. All three potentials satisfy the Helmholtz equation:

$$\nabla^2\widetilde{\phi}_{f,s}+k_{f,s}^2\widetilde{\phi}_{f,s}=0, \quad \nabla^2\widetilde{\boldsymbol{\psi}}+k_t^2\widetilde{\boldsymbol{\psi}}=0, \quad (7)$$

in which the fast, slow, and shear wave numbers are simply $k_{f,s,t}^2=-s^2/c_{f,s,t}^2$, where the phase velocities are given by

$$c_{f,s}=[(\lambda+2N+\alpha A_{f,s})/\bar\rho]^{1/2}, \quad c_t=(N/\bar\rho)^{1/2}. \quad (8)$$

The coefficients $A_f(s)$ and $A_s(s)$ in (6) and (8) are given by

$$A_{f,s}=\frac{D\mp[D^2+(4(\lambda+2N)\alpha^2\bar\rho_{22}^2)/R]^{1/2}}{(2\alpha\bar\rho_{22})/R}, \quad (9)$$

where $D=-[(\lambda+2N)\bar\rho_{22}]/R+\alpha^2\bar\rho_{22}+\bar\rho$. It can be verified that the phase-velocity expressions above agree with the earlier forms given by Biot (1956a, b) and Johnson (1984).

II. ASYMPTOTIC APPROXIMATIONS

To formulate singly and doubly asymptotic approximations, we define at a point on the impedance surface the following fourfold displacement vector (Geers and Chen, 1993):

$$\boldsymbol{u}(\mathbf{x},t)=\{u_1,u_2,u_3,U_n\}^T, \quad (10)$$

where u_1, u_2, and u_3 are orthogonal displacements of the solid phase in a global Cartesian reference frame and U_n is the displacement of the fluid phase normal to the impedance surface. We also define the conjugate traction vector

$$\boldsymbol{t}(\mathbf{x},t)=\{t_1,t_2,t_3,-q\}^T, \quad (11)$$

where t_1, t_2, and t_3 are Cartesian traction components on the solid phase and, again, q is the product of porosity and pore pressure at the impedance surface. Here each component of $\boldsymbol{t}$ has the same direction as its counterpart of $\boldsymbol{u}$.

A. Early-time approximations ETA$_1$ and ETA$_2$

Early-time approximations are obtained through the use of ray theory (Pierce, 1989; Achenbach et al., 1982), as described in Appendix B. The second-order approximation (ETA$_2$) is found to be

$$\boldsymbol{t}(\mathbf{x},t)=\rho c[\boldsymbol{H}(\mathbf{x})\dot{\boldsymbol{u}}(\mathbf{x},t)+\kappa(\mathbf{x})\boldsymbol{C}(\mathbf{x})\boldsymbol{u}(\mathbf{x},t)], \quad (12)$$

in which

$$\rho=\rho_{11}-\rho_{12}^2/\rho_{22},$$

$$c^2=\lim_{s\to\infty}c_f^2=\frac{\Delta+\sqrt{\Delta^2-4(\rho_{11}\rho_{22}-\rho_{12}^2)(PR-Q^2)}}{2(\rho_{11}\rho_{22}-\rho_{12}^2)}, \quad (13)$$

where $\Delta=P\rho_{22}+R\rho_{11}-2Q\rho_{12}$ and $\kappa(\mathbf{x})=(R_\xi+R_\zeta)/(2R_\xi R_\zeta)$ is the local mean curvature of the surface, defined positive when the surface is concave looking into the medium; R_ξ and R_ζ are the two principal radii of curvature. The matrices $\boldsymbol{H}(\mathbf{x})$ and $\boldsymbol{C}(\mathbf{x})$ are given by $\boldsymbol{H}(\mathbf{x})=\boldsymbol{Q}^T(\mathbf{x})\boldsymbol{H}'\boldsymbol{Q}(\mathbf{x})$, and $\boldsymbol{C}(\mathbf{x})=\boldsymbol{Q}^T(\mathbf{x})\boldsymbol{C}'\boldsymbol{Q}(\mathbf{x})$, where $\boldsymbol{Q}(\mathbf{x})$ is a 4×4 transformation matrix with unity as the 4–4 element, zeros as the 1–4, 2–4, 4–1, and 4–2 elements, and with the upper-left 3×3 matrix constituting a rotation of solid–phase displacement

from local to Cartesian global coordinates, and where the local-coordinate matrices H' and C are given by

$$H' = \begin{bmatrix} h_{11} & 0 & 0 & h_{1f} \\ 0 & c_T/c & 0 & 0 \\ 0 & 0 & c_T/c & 0 \\ h_{f1} & 0 & 0 & h_{ff} \end{bmatrix},$$

$$C' = \begin{bmatrix} c_{11} & 0 & 0 & c_{1f} \\ 0 & c_T & 0 & 0 \\ 0 & 0 & c_T & 0 \\ c_{f1} & 0 & 0 & c_{ff} \end{bmatrix}. \tag{14}$$

Expressions for the elements of these matricies appear in Appendix B.

The first-order approximation (ETA$_1$) is found from (12) by taking $\kappa(x)=0$; it is the *plane-wave approximation*. An interesting aspect of (14) is the fact that, in local coordinates, the shear components are uncoupled from each other and from the normal components; the latter two are coupled, however.

The limiting behavior of a pure elastic medium and pure acoustic medium may be recovered from (13) and (14). In the case of the elastic medium, the porosity and the fluid density vanish, which produces $\rho_{11}=\rho_s$, $\rho_{22}=\rho_{12}=0$, and thus $\rho=\rho_s$. Also, $c_f=c_L$, $c_t=c_S$, $c_s=0$, and thus $c=c_L$, where c_L and c_S are the longitudinal- and shear-wave velocities in the elastic medium. These results yield $h_{11}=1$ and $h_{1f}=h_{f1}=h_{ff}=0$.

In the case of the acoustic medium, the porosity becomes unity and the solid density and elastic moduli vanish; this produces $\rho_{22}=\rho_f$, $\rho_{11}=\rho_{12}=0$, $c_f=c_A$, $c_s=c_t=0$, and thus $c=c_A$, where c_A is the acoustic velocity in the fluid. These results yield $h_{ff}=1$ and $h_{11}=h_{1f}=h_{f1}=0$.

B. Late time approximation LTA$_1$

Late-time approximations are obtained by expanding the Laplace-transformed dynamic Green's tensors for the poroelastic medium in powers of s, and then appropriately truncating the series. In Appendix C we truncate at the lowest level, thereby obtaining

$$[B_x + (bs)^{-1}\Delta_x]\, \tilde{t}(\mathbf{x},s) = \Gamma_x \tilde{u}(\mathbf{x},s), \tag{15}$$

in which B_x, Δ_x, and Γ_x are 4×4 matrices of integral operators *in space*: It is convenient to partition these matrices by solid and fluid pertinence as

$$B_x = \begin{bmatrix} B_{ss} & 0 \\ b_{fs} & 0 \end{bmatrix}, \quad \Delta_x = \begin{bmatrix} 0 & \delta_{sf} \\ 0 & \delta_{ff} \end{bmatrix}, \quad \Gamma_x = \begin{bmatrix} \Gamma_{ss} & \gamma_{sf} \\ \Gamma_{fs} & \gamma_{ff} \end{bmatrix}, \tag{16}$$

in which the matrix elements are defined by their action on a surface function $f(\mathbf{x})$ as

$$B_{ss}: \quad B_{ij}\{f(\mathbf{x})\} = \int_S U_{ij}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS',$$

$$b_{fs}: \quad B_{4j}\{f(\mathbf{x})\} = \int_S U_{j4}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS',$$

$$\delta_{sf}: \quad \Delta_{i4}\{f(\mathbf{x})\} = \int_S T_{4i}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS',$$

$$\delta_{ff}: \quad \Delta_{44}\{f(\mathbf{x})\} = \int_S T_{44}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS' + \frac{f(\mathbf{x})}{2},$$

$$\Gamma_{ss}: \quad \Gamma_{ij}\{f(\mathbf{x})\} = \int_S T_{ji}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS' + \frac{\delta_{ij}}{2} f(\mathbf{x}), \tag{17}$$

$$\Gamma_{fs}: \quad \Gamma_{4j}\{f(\mathbf{x})\} = \int_S T_{j4}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS',$$

$$\gamma_{sf}: \quad \Gamma_{i4}\{f(\mathbf{x})\} = \int_S U_{4i}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS',$$

$$\gamma_{ff}: \quad \Gamma_{44}\{f(\mathbf{x})\} = \int_S U_{44}^0(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dS',$$

where $i,j=1,2,3$ and the quasi-static tensors $U_{kl}^0(\mathbf{x},\mathbf{x}')$ and $T_{kl}^0(\mathbf{x},\mathbf{x}')$ for $k,l=1,2,3,4$ are given in Appendix C.

We can manipulate (15) so that $\tilde{t}(\mathbf{x},s)$ appears alone on the left; inverse-transformation of the resulting equation leads to LTA$_1$:

$$t(\mathbf{x},t) = E_x u(\mathbf{x},t) + b F_x \dot{u}(\mathbf{x},t), \tag{18}$$

in which the spatial-operator matrices E_x and F_x are given by

$$E_x = \begin{bmatrix} E_{ss} & e_{sf} \\ 0 & 0 \end{bmatrix},$$

$$F_x = \delta_{ff}^{-1}\begin{bmatrix} 0 & 0 \\ (\Gamma_{fs} - b_{fs}E_{ss}) & (\gamma_{ff} - b_{fs}e_{sf}) \end{bmatrix}, \tag{19}$$

where

$$E_{ss} = (B_{ss} - \delta_{ff}^{-1}\delta_{sf}b_{fs})^{-1}(\Gamma_{ss} - \delta_{ff}\delta_{sf}\Gamma_{fs}),$$

$$e_{sf} = (B_{ss} - \delta_{ff}^{-1}\delta_{sf}b_{fs})^{-1}(\Gamma_{sf} - \delta_{ff}\delta_{sf}\gamma_{ff}). \tag{20}$$

C. Doubly asymptotic approximations DAA$_1$ and DAA$_{1-2}$

Having obtained ETA$_2$ [see (12)], ETA$_1$ [take $\kappa(\mathbf{x})=0$ in (12)] and LTA$_1$ [see (18)], we can obtain two DAAs by the method of operator matching (Nicolas-Vullierme, 1991; Geers and Zhang, 1994; Geers and Lewis, 1997). To obtain DAA$_1$, we seek a unique one-term match of a trial equation to Laplace-transformed ETA$_1$ for $s\to\infty$ and a unique one-term match of the same trial equation to Laplace-transformed LTA$_1$ for $s\to0$. Because both ETA$_1$ and LTA$_1$ contain $\dot{u}(\mathbf{x},t)$ on the right, and because $\rho c H(\mathbf{x}) \neq b F_x$, the most compact trial equation that will produce such matching is

$$(sI + T_0)\, \tilde{t}(\mathbf{x},s) = (s^2 U_2 + s U_1 + U_0)\, \tilde{u}(\mathbf{x},s). \tag{21}$$

But this is equally suitable as a trial equation for a two-term match to Laplace-transformed ETA$_2$ for $s\to\infty$ and a two-term match to Laplace-transformed LTA$_1$ for $s\to0$. Thus we may proceed directly to formulate DAA$_{1-2}$; DAA$_1$ follows immediately as DAA$_{1-2}$ with $\kappa(\mathbf{x})=0$.

The formulation of DAA$_2$ would require both ETA$_2$ and LTA$_2$. The construction of LTA$_2$ involves truncation of the

power-series expansions for the Laplace-transformed Green's tensors at the level immediately above that used to construct LTA$_1$. The results may not be worth the effort, as previous experience in acoustics and elastodynamics has demonstrated that DAA$_{1-2}$ yields high accuracy at low computational cost.

To match (21) to Laplace-transformed (12) for $s \to \infty$, we write (21) as $(I + s^{-1}T_0)\widetilde{t} = (U_2 + s^{-1}U_1 + s^{-2}U_0)s\widetilde{u}$ and preoperate both side of this equation with $(I - s^{-1}T_0)$; this yields

$$[I + O(s^{-2})]\widetilde{t} = [U_2 + s^{-1}(U_1 - T_0 U_2) + O(s^{-2})]s\widetilde{u}. \tag{22}$$

Matching the two leading terms on the right side of this equation to the right side of Laplace-transformed (12), we obtain $U_2 = \rho c H(\mathbf{x})$ and $U_1 - T_0 U_2 = \rho c \kappa(\mathbf{x})C(\mathbf{x})$.

To match (21) to Laplace-transformed (18) for $s \to 0$, we preoperate both sides of (21) with $(T_0 - sI)$, which yields

$$[T_0 T_0 + O(s^2)]\widetilde{t} = [T_0 U_0 + s(T_0 U_1 - U_0) + O(s^2)]\widetilde{u}. \tag{23}$$

Then we preoperate both sides of Laplace-transformed (18) with $T_0 T_0$ and match to (23); this gives $T_0 U_0 = T_0 T_0 E_x$ and $T_0 U_1 - U_0 = b T_0 T_0 F_x$.

Finally, we solve our four equations for the four unknown operators to obtain, with $\Omega_x \equiv [E_x - \rho c \kappa(\mathbf{x})C(x)] \times [\rho c H(\mathbf{x}) - b F_x]^{-1}$,

$$T_0 = \Omega_x, \quad U_2 = \rho c H(\mathbf{x}),$$
$$U_1 = \rho c[\Omega_x H(\mathbf{x}) + \kappa(\mathbf{x})C(\mathbf{x})], \quad U_0 = \Omega_x E_x. \tag{24}$$

Thus introducing these into (21) and inverse-transforming, we obtain DAA$_{1-2}$ as

$$\dot{t}(\mathbf{x},t) + \Omega_x t(\mathbf{x},t) = \rho c H(\mathbf{x})\ddot{u}(\mathbf{x},t) + \rho c[\Omega_x H(\mathbf{x})$$
$$+ \kappa(\mathbf{x})C(\mathbf{x})]\dot{u}(\mathbf{x},t) + \Omega_x E_x u(\mathbf{x},t). \tag{25}$$

We retrieve DAA$_1$ from this equation by taking $\kappa(\mathbf{x}) = 0$.

D. Boundary element discretization

As the operators contained in (25) cannot be analytically evaluated, boundary-element discretization must be employed to obtain solutions to practical problems (Zimmerman and Stern, 1994). To obtain the boundary-element DAA forms, we generate a finite-element mesh on the impedance surface and employ the usual finite-element interpolations

$$u(\mathbf{x},t) \approx \underline{v}(\mathbf{x})\underline{u}(t), \quad t(\mathbf{x},t) \approx \underline{v}(\mathbf{x})\underline{t}(t), \tag{26}$$

where $\underline{v}(\mathbf{x})$ is a row vector of shape functions, and where $\underline{u}(t)$ and $\underline{t}(t)$ are column vectors of nodal displacements and tractions, respectively.

First, we discretize ETA$_2$ by applying Galerkin approximation to (12) (Brebbia et al., 1984), which yields

$$\underline{t}(t) = \rho c[\underline{H}\dot{\underline{u}}(t) + \underline{\kappa C}\underline{u}(t)], \tag{27}$$

in which the matrices $\underline{H}$ and $\underline{\kappa C}$ are given by

$$\underline{H} = \underline{J}^{-1}\int_S \underline{w}H\underline{v}\, dS, \quad \underline{\kappa C} = \underline{J}^{-1}\int_S \underline{w}\kappa C\underline{v}\, dS, \tag{28}$$

where $\underline{J} = \int_S \underline{w}\underline{v}\, dS$. In these equations, $\underline{w}(x)$ is a column vector of chosen weighting functions; it may or may not be the same as the transpose of $\underline{v}(x)$.

Next, we similarly discretize LTA$_1$, as given by (15), to get

$$[\underline{B} + (bs)^{-1}\underline{\Delta}]\widetilde{\underline{t}}(s) = \underline{\Gamma}\, \widetilde{\underline{u}}(s), \tag{29}$$

where

$$\underline{B} = \int_S \underline{w}\mathbf{B}_x\underline{v}\, dS, \quad \underline{\Delta} = \int_S \underline{w}\Delta_x\underline{v}\, dS, \quad \underline{\Gamma} = \int_S \underline{w}\,\Gamma_x\,\underline{v}\, dS. \tag{30}$$

Then we partition these matrices in the manner of (16) and follow the procedure used in Appendix C to proceed from (16) to (18); this yields

$$\underline{t}(t) = \underline{E}\underline{u}(t) + b\underline{F}\dot{\underline{u}}(t), \tag{31}$$

where $\underline{E}$ and $\underline{F}$ are the matrix counterparts of E_x and F_x in (19).

Finally, we introduce the trial equation

$$(s\underline{I} + \underline{T}_0)\widetilde{\underline{t}}(s) = (s^2\underline{U}_2 + s\underline{U}_1 + \underline{U}_0)\widetilde{\underline{u}}(s) \tag{32}$$

and follow the matching procedure of the previous subsection to obtain the boundary-element DAA$_{1-2}$

$$\dot{\underline{t}}(t) + \underline{\Omega}\underline{t}(t) = \rho c \underline{H}\ddot{\underline{u}}(t) + \rho c[\underline{\Omega}\underline{H} + \underline{\kappa C}]\dot{\underline{u}}(t) + \underline{\Omega}\underline{E}\underline{u}(t), \tag{33}$$

where $\underline{\Omega} = [\underline{E} - \rho c \underline{\kappa C}][\rho c \underline{H} - b\underline{F}]^{-1}$. DAA$_1$ follows when $\underline{\kappa C}$ is taken as zero.

III. EVALUATION PROBLEM

We include here an exploratory examination of poroelastodynamic DAA accuracy by considering the problem of a thin, elastic spherical shell embedded in an infinite poroelastic medium and excited by an internal, uniform, step-pressure loading. The impedance surface is placed at the shell-medium interface; thus at this interface, the radial motion of the medium's solid and fluid phases are the same, but the radial loadings are different. The shell is then removed entirely, and the same loading is applied directly to the medium; in this case, the radial surface loadings are the same, but the radial surface motions are different.

A. Formulation

Because of the polar symmetry in this evaluation problem, formulation is greatly simplified. The wave equations for the scalar potentials are given by the first of (7), but the vector potential vanishes. It can be shown that solutions for the two scalar potentials in the Laplace-transform domain are

$$\widetilde{\phi}_f = P_f(s)k_0(rs/c_f), \quad \widetilde{\phi}_s = P_s(s)k_0(rs/c_s), \tag{34}$$

where k_0 is the modified spherical Bessel function of zeroth order, given by

$$k_0(\xi) = \xi^{-1}e^{-\xi}, \tag{35}$$

and P_f and P_s are two arbitrary functions of s. Using (34), (6), and (2), we find the following exact relation, at the impedance surface $r = a$, between the generalized traction vec-

tor $\widetilde{\boldsymbol{t}}=\{\widetilde{t}_n,-\widetilde{q}\}^T$ and the generalized displacement vector $\widetilde{\boldsymbol{u}}=\{\widetilde{u}_n,\widetilde{U}_n\}^T$:

$$\widetilde{\boldsymbol{t}}=\begin{bmatrix} sb_{11}(s)+4N/a & sb_{12}(s) \\ sb_{21}(s) & sb_{22}(s) \end{bmatrix}\widetilde{\boldsymbol{u}}, \qquad (36)$$

where

$$b_{11}(s)=\frac{\bar{\rho}c_f^2c_s^2}{(\lambda+2N)(c_f^2-c_s^2)}\left\{\frac{c_s}{1+c_s/as}\left[\bar{\rho}_{12}\alpha+\bar{\rho}_{11}\right.\right.$$

$$+(\lambda+2N)\left(\frac{\bar{\rho}_{12}^2}{\bar{\rho}_{22}c_s^2\bar{\rho}}-\frac{1}{c_f^2}\right)$$

$$-\frac{\bar{\rho}\bar{\rho}_{12}}{\bar{\rho}_{22}\alpha}\left(1-\frac{\lambda+2N}{\bar{\rho}c_f^2}\right)\left(1-\frac{\lambda+2N}{\bar{\rho}c_s^2}\right)\right]$$

$$-\frac{c_f}{1+c_f/as}\left[(\lambda+2N)\left(\frac{\bar{\rho}_{12}^2}{\bar{\rho}_{22}c_f^2\bar{\rho}}-\frac{1}{c_s^2}\right)\right.$$

$$-\frac{\bar{\rho}\bar{\rho}_{12}}{\bar{\rho}_{22}\alpha}\left(1-\frac{\lambda+2N}{\bar{\rho}c_f^2}\right)\left(1-\frac{\lambda+2N}{\bar{\rho}c_s^2}\right)$$

$$\left.\left.+\bar{\rho}_{12}\alpha+\bar{\rho}_{11}\right]\right\},$$

$$b_{12}(s)=\frac{-\bar{\rho}_{22}c_f^2c_s^2}{(\lambda+2N)(c_f^2-c_s^2)}$$

$$\times\left\{\frac{c_s}{1+c_s/as}\left[-\frac{Q}{R}\bar{\rho}+(\lambda+2N)\frac{\bar{\rho}_{12}}{\bar{\rho}_{22}c_s^2}\right]\right.$$

$$\left.-\frac{c_f}{1+c_f/as}\left[-\frac{Q}{R}\bar{\rho}+(\lambda+2N)\frac{\bar{\rho}_{12}}{\bar{\rho}_{22}c_s^2}\right]\right\},$$

$$\qquad (37)$$

$$b_{21}(s)=\frac{-c_f^2c_s^2}{(\lambda+2N)(c_f^2-c_s^2)\alpha}$$

$$\times\left\{\left[\bar{\rho}-\frac{\lambda+2N}{c_s^2}+\bar{\rho}_{12}\alpha\right]\frac{\rho c_f^2-(\lambda+2N)}{c_f(1+c_f/as)}\right.$$

$$\left.-\left[\bar{\rho}-\frac{\lambda+2N}{c_f^2}+\bar{\rho}_{12}\alpha\right]\frac{\bar{\rho}c_s^2-(\lambda+2N)}{c_s(1+c_s/as)}\right\},$$

$$b_{22}(s)=\frac{-\bar{\rho}_{22}c_f^2c_s^2}{(\lambda+2N)(c_f^2-c_s^2)}\left\{\frac{\bar{\rho}c_f^2-(\lambda+2N)}{c_f(1+c_f/as)}\right.$$

$$\left.-\frac{\bar{\rho}c_s^2-(\lambda+2N)}{c_s(1+c_s/as)}\right\}.$$

Note that most parameters are functions of s.

The corresponding DAA$_{1\text{-}2}$ may be obtained directly from (36) as follows. In the limit of large s (early time), (36) can be reduced to

$$\widetilde{\boldsymbol{t}}=\left\{\begin{bmatrix} b_{11}^\infty & b_{12}^\infty \\ b_{21}^\infty & b_{22}^\infty \end{bmatrix}+\frac{4N}{as}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\}s\widetilde{\boldsymbol{u}}, \quad s\to\infty, \qquad (38)$$

where the constants $b_{ij}^\infty=\lim_{s\to\infty}[b_{ij}(s)]$. Inverse transformation then gives ETA$_2$:

$$\boldsymbol{t}=\begin{bmatrix} b_{11}^\infty & b_{12}^\infty \\ b_{21}^\infty & b_{22}^\infty \end{bmatrix}\dot{\boldsymbol{u}}+\frac{4N}{a}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\boldsymbol{u}. \qquad (39)$$

In the limit of small s (late time), (36) can be reduced to

$$\widetilde{\boldsymbol{t}}=\left\{\frac{4N}{a}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}+sba\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right\}\widetilde{\boldsymbol{u}}, \quad s\to0. \qquad (40)$$

Inverse transformation then gives LTA$_1$,

$$\boldsymbol{t}=\frac{4N}{a}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\boldsymbol{u}+ba\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\dot{\boldsymbol{u}}. \qquad (41)$$

Finally, we identify (39) with (12) and (41) with (18) to formulate DAA$_{1\text{-}2}$ as

$$\dot{\boldsymbol{t}}+\boldsymbol{\Omega}\boldsymbol{t}=\begin{bmatrix} b_{11}^\infty & b_{12}^\infty \\ b_{21}^\infty & b_{22}^\infty \end{bmatrix}\ddot{\boldsymbol{u}}+\left[\boldsymbol{\Omega}\begin{bmatrix} b_{11}^\infty & b_{12}^\infty \\ b_{21}^\infty & b_{22}^\infty \end{bmatrix}-\frac{4N}{a}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right]\dot{\boldsymbol{u}}$$

$$+\boldsymbol{\Omega}\frac{4N}{a}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\boldsymbol{u}, \qquad (42)$$

where

$$\boldsymbol{\Omega}=\frac{8N}{aB}\begin{bmatrix} b_{22}^\infty-ab & -b_{12}^\infty \\ 0 & 0 \end{bmatrix},$$

$$B=b_{11}^\infty(b_{22}^\infty-ab)-b_{12}^\infty(b_{21}^\infty-ab). \qquad (43)$$

The simplicity of these equations relative to (36) and (37) is apparent.

For radial vibration of the thin spherical shell, the governing equation is (Junger and Feit, 1986)

$$\ddot{w}(t)+2(1+v_0)\left(\frac{c_0}{a}\right)^2w=\frac{1}{\rho_0h}(p_0-t_r), \qquad (44)$$

where $w(t)$ is the radial displacement of the middle surface of the shell, $p_0(t)$ is the internal pressure loading on the shell, and $t_r(t)$ is the radial traction at the impedance surface; also, a is the shell's radius, ρ_0 and v_0 are the mass density and Poisson's ratio for the shell material, and $c_0=[E_0/\rho_0(1-v_0^2)]^{1/2}$ is the plate velocity for the shell, where E_0 is Young's modulus for the shell material.

B. Exact and DAA results

For the step-pressure-loaded spherical shell, the radial displacements of the solid and fluid phases at the interface must equal the radial displacement of the shell. Therefore, $w=u_n=U_n$; also $t_r=t_n-q$. After Laplace-transformation, (36) and (44) then yield

$$s^2\widetilde{w}+\left[2(1+v_0)\left(\frac{c_0}{a}\right)^2+\frac{1}{\rho_0h}\left\{\frac{4N}{a}+s[b_{11}(s)+b_{12}(s)\right.\right.$$

$$\left.\left.+b_{21}(s)+b_{22}(s)]\right\}\right]\widetilde{w}=\frac{1}{\rho_0h}\widetilde{p}_0. \qquad (45)$$

For a step loading, $\widetilde{p}_0=P_0/s$, where P_0 is the magnitude of the loading.

For the step-pressure-loaded spherical cavity, the surface loading is the same for both the fluid and solid phases, but
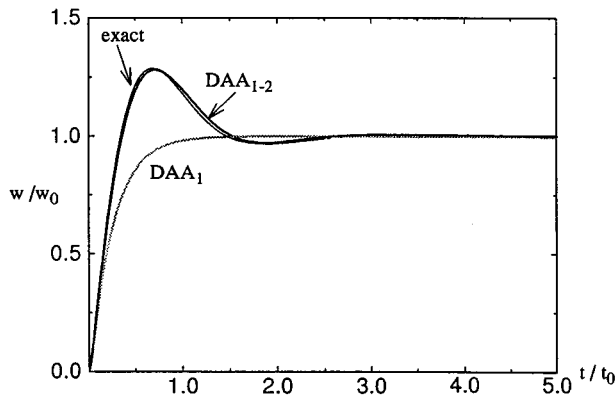
FIG. 1. DAA and exact radial displacement histories for a spherical shell embedded in a poroelastic medium under internal step-pressure loading.

FIG. 2. DAA and exact radial displacement histories for the solid phase at the surface of a spherical cavity under internal step-pressure loading.

displacements in the two phases are independent. The desired relation is obtained by inverting (36), which yields

$$\widetilde{\boldsymbol{u}} = \frac{1}{G(s)} \begin{bmatrix} sb_{22}(s) & -sb_{12}(s) \\ -sb_{21}(s) & sb_{11}(s) + 4N/a \end{bmatrix} \widetilde{\boldsymbol{t}}, \qquad (46)$$

where $G(s) = (4N/a)sb_{22}(s) + s^2[b_{11}(s)b_{22}(s) - b_{12}(s)b_{21}(s)]$. Because $\widetilde{q} = f\widetilde{p}$ and $\widetilde{p} = \widetilde{t}_n = \widetilde{p}_0$, we find

$$\widetilde{u}_1 = \frac{s}{G(s)} [b_{22}(s) - fb_{12}(s)]\widetilde{p}_0$$

and $\hspace{9cm}$ (47)

$$\widetilde{U}_n = \frac{1}{G(s)} \{-sb_{21}(s) + f[sb_{11}(s) + 4N/a]\}\widetilde{p}_0.$$

"Exact" time-domain results may be obtained by inverting (45) and (47) numerically. Here we use the Stehfest scheme (Davies and Martin, 1979) to generate results for the following sets of parameters. For the Biot medium,

$\rho_{11} = 1405.5$ kg/m^3, $\rho_{12} = -0.25$ kg/m^3,

$\rho_{22} = 470$ kg/m^3, $b = 2.21 \times 10^6$ N s/m^4

$N = 2.61 \times 10^7$ Pa, $P = 1.546 \times 10^9$ Pa,

$Q = 9.76 \times 10^8$ Pa, $R = 8.858 \times 10^8$ Pa

which, taken from Zimmerman and Stern (1993), represent the properties of coarse sand sediment. For the shell,

$a = 1.0$ m, $h = 0.01$ m, $\rho_0 = 7700$ kg/m^3,

$E_0 = 1.195 \times 10^{11}$ Pa, $\nu_0 = 0.28$,

the last three of which represent the properties of steel.

Transform inversion is not necessary to obtain DAA-based solutions. For the spherical shell, one merely solves (44) and (42) simultaneously with $t_r = t_n - q$, $\mathbf{t} = \{t_n - q_n\}^T$, and $\mathbf{u} = \{w \quad w\}^T$ using any reliable time-integration algorithm. For the spherical cavity, one simply solves (42) for $\mathbf{u} = \{u_n \quad U_n\}^T$ with $\mathbf{t} = \{p_0 \quad fp_0\}^T$.

Figure 1 shows exact, DAA$_{1-2}$ and DAA$_1$ displacement histories for the step-loaded spherical shell; $w_0 = P_0/[2\rho_0 h(1-\nu_0)(c_0/a)^2 + 4N/a]$ is the static radial displacement under constant pressure P_0 and $t_0 = a/c$. We see

that DAA$_1$ yields accurate results at early and late time, but fails to produce the pronounced overshoot. In contrast, DAA$_{1-2}$ yields an accurate response for all time. The response behavior exhibited in Fig. 1 is quite similar to that displayed by a spherical shell embedded in an elastic medium (Geers *et al.*, 1997). A more comprehensive examination is required, however, to draw definitive conclusions in this regard.

Exact and DAA$_{1-2}$ results for the step-loaded spherical cavity are shown in Figs. 2 and 3 for the solid and fluid phases, respectively. The response behavior exhibited in Fig. 2 is quite different from that displayed by a spherical cavity in an elastic medium (Geers and Lewis, 1997). In fact, the latter more closely resembles the behavior shown in Fig. 1. In Fig. 2, $u_0 = aP_0/4N$, which is the static radial displacement under constant pressure P_0. In Fig. 3, the fluid displacement increases without bound at late time. We observe good agreement between the exact and DAA$_{1-2}$ histories in both figures.

IV. CONCLUSION

Singly and doubly asymptotic approximations have been formulated for a poroelastic medium described by Biot's equations. The early-time approximation ETA$_2$ was obtained
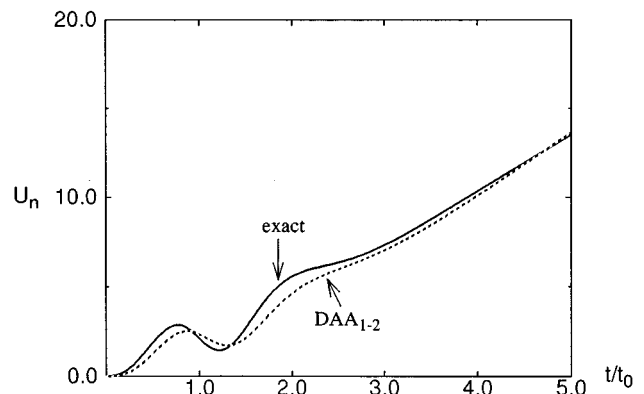
FIG. 3. DAA and exact radial displacement histories for the fluid phase at the surface of a spherical cavity under internal step-pressure loading.

with ray theory (Geers and Lewis, 1997). The late-time approximation LTA$_1$ was found by using a modified version of a generalized reciprocal work relation (Zimmerman and Stern, 1993) and expanding the corresponding Green's functions in the limit of small s. The doubly asymptotic approximations DAA$_1$ and DAA$_{1-2}$ were then formulated by the method of operator matching (Nicolas-Vullierme, 1991). All approximations obtained reduce to their counterparts for purely elastic and purely acoustic domains. Matrix poroelastic DAAs were given for boundary-element analysis of complex surfaces. A preliminary evaluation of DAA accuracy was carried out by examining the surface response of a step-pressurized spherical shell and spherical cavity located in an infinite poroelastic domain; good agreement between DAA$_{1-2}$ and exact displacement histories was observed.

ACKNOWLEDGMENTS

APPENDIX A: MODE DECOMPOSITION AND RELATIONS BETWEEN BIOT AND PHYSICAL PARAMETERS

Introducing (6) into (3) and (5) with $\widetilde{F}=0$, $\widetilde{V}=0$, we obtain

$$(a_f \nabla^2 - \rho s^2)\widetilde{\phi}_f + (a_s \nabla^2 - \rho s^2)\widetilde{\phi}_s = 0,$$
$$(A_f \nabla^2 + d_f)\widetilde{\phi}_f + (A_s \nabla^2 + d_s)\widetilde{\phi}_s = 0, \qquad (A1)$$
$$\nabla^2 \widetilde{\boldsymbol{\psi}} + k_t^2 \widetilde{\boldsymbol{\psi}} = 0,$$

where, with f and s denoting "fast" and "slow,"

$$a_{f,s} = \lambda + 2N + \alpha A_{f,s},$$
$$d_{f,s} = -A_{f,s}(\rho_{22}/R)s^2 + \alpha \rho_{22} s^2. \qquad (A2)$$

Compatibility of the first and second of (A1) leads to the formula for $A_{f,s}$ given in (9). We observe that all the equations in (A1) are satisfied when the potentials satisfy the Helmholtz equation, (7). Consequently, in the absence of body forces, three modes of wave propagation are present in the Biot medium: fast wave, slow wave, and shear wave. c_f, c_s, and c_t are the fast-, slow-, and shear-wave speeds, respectively.

Relations between Biot moduli and more physically based moduli are (Biot and Willis, 1957; Johnson, 1984)

$$P = \frac{(1-f)(1-f-\kappa_b/\kappa_s)\kappa_s + f\kappa_b\kappa_s/\kappa_f}{1-f-\kappa_b/\kappa_s + (\kappa_s/\kappa_f)f} + \frac{4}{3}N,$$

$$Q = \frac{(1-f-\kappa_b/\kappa_s)f\kappa_s}{1-f-\kappa_b/\kappa_s + (\kappa_s/\kappa_f)f}, \qquad (A3)$$

$$R = \frac{f^2\kappa_s}{1-f-\kappa_b/\kappa_s + (\kappa_s/\kappa_f)f},$$

where κ_b, κ_s, and N are the solid's "jacketed" bulk modulus, "skeleton" bulk modulus, and shear modulus, respectively, and where κ_f is the fluid's bulk modulus. Biot densities are related to the densities of the fluid and solid phases as

$$\bar{\rho}_{11} = (1-f)\rho_s + (\bar{\alpha}-1)f\rho_f, \quad \bar{\rho}_{22} = f\rho_f + (\bar{\alpha}-1)f\rho_f,$$
$$\bar{\rho}_{12} = -(\bar{\alpha}-1)f\rho_f \qquad (A4)$$

in which ρ_s is the solid's mass density, ρ_f is the fluid's mass density, and $\bar{\alpha} = \alpha + b/(sf\rho_f)$ is the complex tortuosity; the conventional tortuosity α is a purely geometrical quantity. Each density $\bar{\rho}_{ij}$ reduces to ρ_{ij} when $b=0$.

In the limit $f \to 0$, $\rho_f \to 0$, the poroelastic medium reduces to an elastic medium; in the limit $f \to 1$, $N \to 0$, $\kappa_b \to 0$, it reduces to an acoustic fluid. It can be shown that these limits are, in fact, produced by (A3), (A4), (8), and $\lambda + 2N = P - Q^2/R$ with

$$c_f^2 \to \frac{\kappa_s + 4N/3}{\rho_s}, \quad c_s^2 \to 0, \quad c_t^2 \to \frac{N}{\rho_s} \text{ as } f \to 0, \quad \rho_f \to 0, \tag{A5}$$

$$c_f^2 \to \frac{\kappa_f}{\rho_f}, \quad c_s^2 \to 0, \quad c_t^2 \to 0 \text{ as } f \to 1, \quad \kappa_b \to 0, \quad N \to 0.$$

Note that the slow-wave speed vanishes in both limits.

APPENDIX B: EARLY-TIME APPROXIMATION ETA$_2$

Early-time asymptotic relations are obtained in the limit $s \to \infty$. In this limit, the potentials in (7) may be found from ray theory as (Geers and Lewis, 1997)

$$\begin{bmatrix} \widetilde{\phi}_f(x,s) \\ \widetilde{\phi}_s(x,s) \\ \widetilde{\boldsymbol{\psi}}(x,s) \end{bmatrix}$$

$$= \left[\frac{R_\xi R_\zeta}{(\eta+R_\xi)(\eta+R_\zeta)}\right]^{1/2} \begin{bmatrix} \exp\left(-\frac{s\eta}{c_f}\right)\phi_f(\mathbf{x}_0,s) \\ \exp\left(-\frac{s\eta}{c_s}\right)\phi_s(\mathbf{x}_0,s) \\ \exp\left(-\frac{s\eta}{c_t}\right)\psi(\mathbf{x}_0,s) \end{bmatrix}. \quad (B1)$$

Here (ξ,ζ,η) form a triad of local orthogonal coordinates (with η constituting the ray direction), $\mathbf{x}_0$ denotes the position vector for a point on the surface of interest, and R_ξ and R_ζ are the principal radii of surface curvature at $\mathbf{x}_0$. From (B1), we can obtain

$$\frac{\partial}{\partial \eta}\begin{bmatrix} \widetilde{\phi}_f(\mathbf{x},s) \\ \widetilde{\phi}_s(\mathbf{x},s) \\ \widetilde{\boldsymbol{\psi}}(\mathbf{x},s) \end{bmatrix}\Bigg|_{\mathbf{x}_0} = -\begin{bmatrix} \left(\frac{s}{c_f}+\kappa\right)\widetilde{\phi}_f(\mathbf{x}_0,s) \\ \left(\frac{s}{c_s}+\kappa\right)\widetilde{\phi}_s(\mathbf{x}_0,s) \\ \left(\frac{s}{c_t}+\kappa\right)\widetilde{\boldsymbol{\psi}}(\mathbf{x}_0,s) \end{bmatrix}$$

and $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (B2)$

$$\frac{\partial}{\partial \xi}(\ ) = \frac{\partial}{\partial \zeta}(\ ) = 0,$$

where κ is the mean curvature, given by $\kappa \equiv (R_\xi + R_\zeta)/(2R_\xi R_\zeta)$.

Introducing (B1) into (6) we find

$$\widetilde{u}_\eta = \left(\frac{\partial \widetilde{\phi}_s}{\partial \eta} + \frac{\partial \widetilde{\phi}_f}{\partial \eta} \right),$$

$$\widetilde{u}_\xi = -\left(\frac{\partial^2 \widetilde{\psi}_\xi}{\partial \eta^2} + 2\kappa \frac{\partial \widetilde{\psi}_\xi}{\partial \eta} + R_\xi^{-1} R_\zeta^{-1} \widetilde{\psi}_\xi \right),$$

$$\widetilde{u}_\zeta = -\left(\frac{\partial^2 \widetilde{\psi}_\zeta}{\partial \eta^2} + 2\kappa \frac{\partial \widetilde{\psi}_\zeta}{\partial \eta} + R_\xi^{-1} R_\zeta^{-1} \widetilde{\psi}_\zeta \right),$$

$$\widetilde{q} = A_s \nabla^2 \widetilde{\phi}_s + A_f \nabla^2 \widetilde{\phi}_f,$$

(B3)

where $\nabla^2 = \partial^2/\partial \eta^2 + 2\kappa(\partial/\partial \eta)$. Further, from (2) and (B3) we may express the local stress field in terms of the potentials, from which the surface tractions follow as

$$\widetilde{t}_\xi - \frac{Q}{R} \widetilde{q} = 4\kappa N \frac{\partial}{\partial \eta} (\widetilde{\phi}_f + \widetilde{\phi}_s) - (\lambda + 2N)(\nabla^2 \widetilde{\phi}_f)$$
$$- (\lambda + 2N)\nabla^2 \widetilde{\phi}_s,$$

$$\widetilde{t}_\xi = N \frac{\partial}{\partial \eta} \left(\frac{\partial^2 \widetilde{\psi}_\xi}{\partial \eta^2} + 2\kappa \frac{\partial \widetilde{\psi}_\xi}{\partial \eta} + R_\xi^{-1} R_\zeta^{-1} \widetilde{\psi}_\xi \right),$$

(B4)

$$\widetilde{t}_\eta = N \frac{\partial}{\partial \eta} \left(\frac{\partial^2 \widetilde{\psi}_\zeta}{\partial \eta^2} + 2\kappa \frac{\partial \widetilde{\psi}_\zeta}{\partial \eta} + R_\xi^{-1} R_\zeta^{-1} \widetilde{\psi}_\zeta \right).$$

Laplace-transformation of the second of (1) yields for normal fluid displacement,

$$\widetilde{\mathbf{U}} = -\frac{\rho_{12}}{\rho_{22}} \widetilde{\mathbf{u}} - \frac{\nabla(\widetilde{q})}{\rho_{22} s^2}.$$

(B5)

Utilization of (B3) then gives

$$\widetilde{U}_\eta + \frac{\rho_{12}}{\rho_{22}} \widetilde{u}_\eta = -\frac{1}{\rho_{22} s^2} \frac{\partial}{\partial \eta} (A_s \nabla^2 \phi_s + A_f \nabla^2 \phi_f).$$

(B6)

Finally, with (B1), we may express the wave equations of (7) as

$$\frac{\partial^2 \widetilde{\phi}_{f,s}}{\partial n^2} + 2\kappa \frac{\partial \widetilde{\phi}_{f,s}}{\partial \eta} = \frac{s^2}{c_{f,s}^2} \widetilde{\phi}_{f,s},$$

$$\frac{\partial^2 \widetilde{\psi}_\xi}{\partial \eta^2} + 2\kappa \frac{\partial \widetilde{\psi}_\xi}{\partial \eta} + (R_\xi - R_\zeta) R_\zeta^{-1} R_\xi^{-2} \widetilde{\psi}_\xi = \frac{s^2}{c_t^2} \widetilde{\psi}_\xi,$$

(B7)

$$\frac{\partial^2 \widetilde{\psi}_\zeta}{\partial \eta^2} + 2\kappa \frac{\partial \widetilde{\psi}_\zeta}{\partial n} + (R_\zeta - R_\xi) R_\xi^{-1} R_\zeta^{-2} \widetilde{\psi}_\zeta = \frac{s^2}{c_t^2} \widetilde{\psi}_\zeta.$$

Now, eliminating the potentials contained in (B3)–(B7), with help from (B2), we obtain the relation between traction and displacement vectors, given by (12) with t, u, H, and C replaced by t', u', H', and C'. The components in (14) that are produced by this process are

$$h_{11} = \frac{c_{f\infty} c_{s\infty}^2}{(\lambda + 2N)(c_{s\infty} + c_{f\infty})} \left\{ \rho_{11} + \rho_{12} \frac{R\rho_{12} - Q\rho_{22}}{\rho_{12} R} - (\lambda + 2N)\left(\frac{c_{f\infty}^2 + c_{f\infty} c_{s\infty} + c_{s\infty}^2}{c_{f\infty}^2 c_{s\infty}^2} \right) - \frac{\rho_{12}^2}{\rho_{22}} \left(\frac{\lambda + 2N}{\rho_\infty c_{f\infty} c_{s\infty}} \right) \right.$$
$$\left. - \frac{\rho_{12} \rho_\infty R}{(R\rho_{12} - Q\rho_{22}) c_{f\infty}^2 c_{s\infty}^2} \left(c_{s\infty}^2 - \frac{\lambda + 2N}{\rho_\infty} \right)\left(c_{f\infty}^2 - \frac{\lambda + 2N}{\rho_\infty} \right) \right\},$$

$$h_{1f} = \frac{\rho_{22} c_{f\infty} c_{s\infty}^2}{(\lambda + 2N)(c_{s\infty} + c_{f\infty})} \left\{ \frac{R\rho_{12} - Q\rho_{22}}{\rho_{12} R} - \left(1 + \frac{\lambda + 2N}{\rho_\infty c_{f\infty} c_{s\infty}} \right) \frac{\rho_{12} R}{R\rho_{12} - Q\rho_{22}} \right\},$$

$$h_{f1} = \frac{\rho_{12}}{c_{f\infty}} \left[\frac{c_{s\infty}}{c_{f\infty} + c_{s\infty}} + \frac{\rho_\infty c_{f\infty} c_{s\infty}^2}{(\lambda + 2N)(c_{s\infty} + c_{f\infty})} \right] + \frac{\rho_\infty \rho_{22} R(c_{s\infty}^2 - (\lambda + 2N)/\rho_\infty)(c_{f\infty}^2 - (\lambda + 2N)/\rho_\infty)}{(\lambda + 2N)(R\rho_{12} - Q\rho_{22}) c_{f\infty}(c_{f\infty} + c_{s\infty})},$$

(B8)

$$h_{ff} = \frac{\rho_{22}}{\rho_\infty} \left[\frac{c_{s\infty}}{c_{f\infty} + c_{s\infty}} + \frac{\rho_\infty c_{f\infty} c_{s\infty}^2}{(\lambda + 2N)(c_{s\infty} + c_{f\infty})} \right],$$

$$c_{11} = \frac{4N}{\rho_\infty c_{f\infty}} - c_{s\infty} \left[\frac{Q}{R} \frac{\rho_{12} c_{f\infty}^2 c_{s\infty}^2}{\lambda + 2N} + \frac{Q}{R} \frac{\rho_\infty c_{f\infty} c_{s\infty}}{\alpha_\infty(\lambda + 2N)} \left(1 - \frac{\lambda + 2N}{\rho_\infty c_{f\infty}^2} \right)\left(1 - \frac{\lambda + 2N}{\rho_\infty c_{s\infty}^2} \right) \right],$$

$$c_t = \sqrt{N/\rho}, \quad c_{1f} = -\rho_{22} \frac{Q}{R} \frac{c_{f\infty} c_{s\infty}^2}{\lambda + 2N}, \quad c_{ff} = -\frac{\rho_{22}^2 c_{f\infty}^2 c_{s\infty}^3}{(\lambda + 2N)^2},$$

$$c_{f1} = -\frac{\rho_\infty c_{f\infty} c_{s\infty}^2(1 - (\lambda + 2N)/\rho_\infty c_{f\infty}^2)(1 - (\lambda + 2N)/\rho_\infty c_{s\infty}^2)}{(\lambda + 2N)\alpha_\infty} - \frac{\rho_{12} \rho_{22} c_{f\infty} c_{s\infty}^3}{(\lambda + 2N)^2},$$

1368   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Q. Qi and T. L. Geers: Transient poroelastodynamics   1368

where the subscript ∞ denotes the limiting value of the corresponding parameter as $s \to \infty$; from the second of (13), $c_{f\infty} = \lim_{s\to\infty}(c_f) = c$.

We obtain (12) itself by transforming from local to global coordinates using the standard relations

$$\boldsymbol{u}'(\boldsymbol{\xi},t) = \boldsymbol{Q}\boldsymbol{u}(\mathbf{x},t), \quad \boldsymbol{t}'(\boldsymbol{\xi},t) = \boldsymbol{Q}\boldsymbol{t}(\mathbf{x},t), \tag{B9}$$

where $\boldsymbol{u}' = \{u_\xi, u_\zeta, u_\eta, U_n\}^T$, $\boldsymbol{t}' = \{t_\xi, t_\zeta, t_\eta, -q_n\}^T$, $\boldsymbol{u}(\mathbf{x},t) = \{u_1, u_2, u_3, U_n\}^T$ and $\boldsymbol{t} = \{t_1, t_2, t_3, -q\}^T$.

APPENDIX C: LATE TIME APPROXIMATION LTA$_1$

Late-time asymptotic relations are obtained in the limit $s \to 0$. Appropriate Green's-function and reciprocal-work relations have been obtained by Zimmerman and Stern (1993) in the frequency domain; the following corresponds to their formulation expressed in the Laplace-transform domain.

We define two vectors $\mathbf{W}$ and $\mathbf{G}$ as

$$\mathbf{W} \equiv \{w_1, w_2, w_3, w_4\}^T = \{\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{q}\}^T,$$

$$\mathbf{G} \equiv \{g_1, g_2, g_3, g_4\}^T = \{\tilde{F}_1, \tilde{F}_2, \tilde{F}_3, \tilde{V}\}^T. \tag{C1}$$

Then, from Zimmerman and Stern (1993),

$$w_i(\mathbf{x}) = \int_S [U_{ji}(\mathbf{x},\mathbf{y})t_j(\mathbf{y}) - T_{ji}(\mathbf{x},\mathbf{y})w_j(\mathbf{y})]dS(\mathbf{y}),$$

$$i = 1,2,3,4, \tag{C2}$$

where S denotes a closed surface, Ω denotes the surrounding porous medium, and U_{ji} and T_{ji} are Cartesian components of the Green's displacement and stress tensors.

The U_{ji} are given by

$$U_{ji} = \frac{\partial^2}{\partial y_j \, \partial y_i}(\phi_{f1} + \phi_{s1}) + \delta_{ji}\frac{\partial^2 \psi_1}{\partial y_k \, \partial y_k} - \frac{\partial^2 \psi_1}{\partial y_j \, \partial y_i},$$

$$i,j = 1,2,3,$$

$$U_{i4} = \frac{\partial}{\partial y_i}(\phi_{f2} + \phi_{s2}), \quad i = 1,2,3,$$

$$U_{4i} = \frac{\partial}{\partial y_i}\frac{\partial^2}{\partial y_k \, \partial y_k}(A_f\phi_{f1} + A_s\phi_{s1}), \quad i = 1,2,3, \tag{C3}$$

$$U_{44} = \frac{\partial^2}{\partial y_k \, \partial y_k}(A_f\phi_{f2} + A_s\phi_{s2}),$$

where ϕ_{f1}, ϕ_{f2}, ϕ_{s1}, and ϕ_{s2} are the Green's-function potentials; expressions for these are given below. In (C3), the origin of the reference coordinate system coincides with the observation point, i.e., $r = |\bar{y}|$. Note that our expressions for A_f and A_s are different from those of Zimmerman and Stern (see Sec. I B). The T_{ji} are given by

$$T_{ji} = \lambda U_{ki,k}\, n_j + N(U_{ij,k}\, n_k + U_{ki,j}\, n_k) - \frac{Q}{R}\, U_{4i}n_j,$$

$$i = 1,2,3,4; \quad j = 1,2,3, \tag{C4}$$

$$T_{4i} = -\left(\frac{\bar{\rho}_{12}}{\bar{\rho}_{22}}\, U_{ki}\, n_k + \frac{U_{44,i}\, n_j}{\bar{\rho}_{22}s^2}\right), \quad i = 1,2,3,4.$$

The Green's-function potentials in (C3) are

$$\phi_{f1} = \frac{C_{1f}}{r}(1 - e^{ik_f r}), \quad \phi_{s1} = \frac{C_{1s}}{r}(1 - e^{ik_s r}),$$

$$\psi_1 = \frac{C_t}{r}(1 - e^{ik_t r}), \tag{C5}$$

$$\phi_{f2} = \frac{C_{2f}}{r}(1 - e^{ik_f r}), \quad \phi_{s2} = \frac{C_{2s}}{r}(1 - e^{ik_s r}),$$

where

$$C_{1f} = -\frac{\alpha R - A_s}{4\pi\rho(A_f - A_s)}, \quad C_{1s} = -\frac{\alpha R - A_f}{4\pi\rho(A_s - A_f)},$$

$$C_t = \frac{1}{4\pi N k_t^2}, \tag{C6}$$

$$C_{2f} = \frac{R}{4\pi s^2(A_f - A_s)\rho_{22}} = -C_{2s}.$$

The wave numbers are given by $k_{f,s,t} = is/c_{f,s,t}$; other parameters are defined in (9). It can be shown from (A1) that, when $c_s^2 = a_s/\rho \to 0$, the appropriate forms of ϕ_{s1} and ϕ_{s2} are $\phi_{s1,2} = C_{1,2s}/r$. This corresponds to $k_s = is/c_s \to i\infty$ in (C5).

Low-frequency behavior is found by expansion of the potential functions of (C5) in the limit $s \to 0$. Taking notice that most parameters are s dependent, we obtain

$$\phi_{f1} = D_{f1}\left(\frac{1}{c_{f0}s} - \frac{r}{2c_{f0}^2}\right), \quad \phi_{s1} = D_{s1}\left(\frac{1}{s^{1/2}c_{s0}} - \frac{r}{2c_{s0}^2}\right),$$

$$\psi_1 = D_t\left(\frac{r}{2c_{t0}^2} - \frac{1}{c_{t0}s^{1/2}}\right), \tag{C7}$$

$$\phi_{f2} = D_{f2}\frac{1}{c_{f0}}, \quad \phi_{s2} = D_{f2}\left(\frac{r}{2c_{s0}^2} - \frac{1}{s^{1/2}c_{s0}}\right),$$

in which

$$c_{f0}^2 = \frac{P + R + 2Q}{\rho^0}, \quad c_{s0}^2 = \frac{PR - Q^2}{b(P + R + 2Q)}, \quad c_t^2 = \frac{N}{\rho^0},$$

$$D_{f1} = -\frac{1}{4\pi\rho^0}, \quad D_{s1} = \frac{PR - Q^2}{4\pi b(P + R + 2Q)^2}, \tag{C8}$$

$$D_t = -D_{f1}, \quad D_{f2} = \frac{R + Q}{4\pi b(P + R + 2Q)},$$

where $\rho^0 = \rho_{11} + \rho_{22} + 2\rho_{12}$.

The introduction of (C7) into (C3) yields the Green's displacement tensors for the Biot medium in the low-frequency limit:

$$U^0_{ji} = \frac{\delta_{ji}}{2r}\left[\frac{D_t}{c^2_{t0}} - \left(\frac{D_{f1}}{c^2_{f0}} + \frac{D_{s1}}{c^2_{s0}}\right)\right]$$

$$+ \frac{y_i y_j}{2r^3}\left[\frac{D_t}{c^2_{t0}} + \left(\frac{D_{f1}}{c^2_{f0}} + \frac{D_{s1}}{c^2_{s0}}\right)\right], \quad i,j=1,2,3,$$

$$U^0_{i4} = \frac{D_{f2}}{2c^2_{s0}}\frac{y_i}{r},$$

$$\hspace{3cm}(C9)$$

$$U^0_{4i} = \left(\frac{D_{f1}}{c^2_{f0}}A_{f0} + \frac{D_{s1}}{c^2_{s0}}A_{s0}\right)\frac{y_i}{r^3}, \quad i=1,2,3,4,$$

$$U^0_{44} = \frac{D_{f2}}{r}\frac{A_{s0}}{c^2_{s0}},$$

where

$$A_{f0} = \frac{PR - Q^2}{R+Q}, \quad A_{s0} = -(Q+R). \hspace{1cm}(C10)$$

From these, the Green's traction tensors may be obtained as, with $i,j = 1,2,3,$

$$T^0_{ji} = N\left\{\Delta_1\left[\frac{n_i y_j + n_j y_i + \delta_{ij}y_k n_k}{r^3} - \frac{3y_i y_j y_k n_k}{r^5}\right]\right.$$

$$\left. - \frac{D_t}{c^2_{t0}}\left(\frac{n_i y_j}{r^3} + \frac{\delta_{ij}y_k n_k}{r^3}\right)\right\} + \lambda\left(\Delta_1 - \frac{D_t}{c^2_{t0}}\right)\frac{y_i n_j}{r^3}$$

$$- \frac{Q}{R}\left(A_{f0}\frac{D_{f1}}{c^2_{f0}} + A_{s0}\frac{D_{s1}}{c^2_{s0}}\right)\frac{y_i n_j}{r^3},$$

$$T^0_{j4} = N\left\{\left(\frac{n_j}{r^3} - \frac{3y_j y_k n_k}{r^5}\right)\left[\left(A_{f0}\frac{D_{f1}}{c^2_{f0}} + A_{s0}\frac{D_{s1}}{c^2_{s0}}\right) + \frac{D_{s2}}{2c^2_{s0}}\right]\right\}$$

$$+ \left(\lambda\frac{D_{f2}}{c^2_{s0}} - \frac{Q}{R}\frac{D_{f2}}{c^2_{s0}}A_{s0}\right)\frac{n_j}{r},$$

$$\hspace{3cm}(C11)$$

$$T^0_{4j} = \frac{1}{(\bar{\rho}_{22}s^2)}\left(\frac{D_{f1}}{c^2_{f0}}A_{f0} + \frac{D_{s1}}{c^2_{s0}}A_{s0}\right)\left(\frac{n_j}{r^3} - \frac{3y_j y_k n_k}{r^5}\right)$$

$$- \frac{n_j}{2r}\Delta_2 + \frac{y_j y_k n_k}{2r^3}\Delta_2,$$

$$T^0_{44} = \frac{D_{f2}}{2c^2_{s0}}\frac{y_k n_k}{r} - \frac{D_{f2}}{(\bar{\rho}_{22}s^2)}\frac{A_{s0}}{c^2_{s0}}\frac{y_k n_k}{r^3},$$

where

$$\Delta_1 = \left(\frac{D_{f1}}{c^2_{f0}} + \frac{D_t}{c^2_{t0}} + \frac{D_{s1}}{c^2_{s0}}\right), \quad \Delta_2 = \left(-\frac{D_{f1}}{c^2_{f0}} + \frac{D_t}{c^2_{t0}} - \frac{D_{s1}}{c^2_{s0}}\right).$$

$$\hspace{3cm}(C12)$$

Now that all the low-frequency Green's-tensor components are available, we may proceed to derive the late-time asymptotic approximations. After some algebra, we find from (C2)

$$\begin{bmatrix} \mathbf{B_{ss}} & \dfrac{\delta_{sf}}{bs} \\[2mm] \mathbf{b_{fx}} & \dfrac{\delta_{ff}}{bs} \end{bmatrix}\tilde{t} = \begin{bmatrix} \mathbf{\Gamma_{ss}} & \boldsymbol{\gamma}_{sf} \\ \mathbf{\Gamma}_{fs} & \gamma_{ff} \end{bmatrix}\tilde{\mathbf{u}}, \hspace{1cm}(C13)$$

where the operator components are defined in (17). This equation can be rearranged and inverse-transformed to obtain (18).

In the reminder of this Appendix, we examine the limiting cases of a pure elastic medium and a pure acoustic medium. One common feature of these limits is the vanishing value of the slow wave speed c_s (Appendix A).

1. Pure elastic medium: $f \to 0, \rho_f \to 0$

In this limit, $e^{ik_s r}$ disappears from both ϕ_{s1} and ϕ_{s2} in (C5). Furthermore, according to (A3), $Q \to 0$, $R \to 0$, and $P \to \kappa_s + 4N/3$. The interaction terms in (C13) vanish with the fluid phase, and $N/P \to 1/(4/3 + k_s/N) = (1-2\nu)/2(1-\nu)$. Thus the components of the Green's displacement tensors for the poroelastic medium reduce to

$$U^0_{ji} = \frac{1}{16\pi N}\left[(3-4\nu)\frac{\delta_{ji}}{r} + \frac{y_j y_i}{r^3}\right], \hspace{1cm}(C14)$$

which we recognize as those of Kelvin's solution for an elastic medium. The same limiting process yields from (C4) the components of the Green's stress tensor for an elastic medium:

$$T^0_{ji} = \frac{-1}{8\pi(1-\nu)}\left\{\frac{y_k n_k}{r^3}\left[(1-2\nu)\delta_{ji} + \frac{3y_j y_i}{r^2}\right]\right.$$

$$\left. - (1-2\nu)\left(\frac{y_i n_j - y_j n_i}{r^3}\right)\right\}. \hspace{1cm}(C15)$$

Finally, if we restrict the indices i,j to vary from 1 to 3, then (C2), together with the Green's tensors of (C14) and (C15) reduce to the familiar Somigliana' identity (Cruse and Rizzo, 1968).

2. Pure acoustic medium: $f \to 1$, $\kappa_b \to 0$, $N \to 0$

In this limit, $e^{ik_s r}$ again disappears from both ϕ_{s1} and ϕ_{s2} in (C5). Furthermore, according to (A3), $Q \to 0$, $P \to 0$, $N \to 0$, and $R \to \kappa_f$. The interaction terms in (C13) vanish with the solid phase and so does the drag parameter b. Careful evaluation of the second matrix equation of (C13) in this limit results in

$$-4\pi C\tilde{p} + \int_S\left(\frac{y_k n_k}{r^3}\right)\tilde{p}\,dS = \rho_f\int_S\frac{\tilde{a}_f}{r}\,dS, \hspace{1cm}(C16)$$

where ρ_f is the fluid density and $\tilde{a}_f = s^2\tilde{U}_f$ is the fluid acceleration. The minus sign in front of $4\pi C\tilde{p}$ may be eliminated by changing the signs of the normal of the surface and the displacement so that they point outward for the closed surface S. Finally, when the observation point lies on the surface, $4\pi C\tilde{p}$ may be incorporated in the integral on the left to convert the Cauchy principle integral to a regular integral. This yields, with $n'_k = -n_k$ and $\tilde{a}'_f = -\tilde{a}_f$,

$$\int_{\partial\Omega} \left(\frac{y_k n_k'}{r^3} \right) \widetilde{p} \, dS = \rho_f \int_{\partial\Omega} \frac{\widetilde{a}_f'}{r} \, dS, \qquad (C17)$$

which is the well-known hydrodynamic relation.

Achenbach, J. D., Gautesen, A. K., and McMaken, H. (**1982**). *Ray Methods for Waves in Elastic Solids* (Pitman, London).

Biot, M. A. (**1956a**). ''Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low frequency range,'' J. Acoust. Soc. Am. **28**, 168–178.

Biot, M. A. (**1956b**). ''Theory of propagation of elastic waves in a fluid-saturated porous solid. II. High frequency range,'' J. Acoust. Soc. Am. **28**, 179–191.

Biot, M. A., and Willis, D. G. (**1957**). ''The elastic coefficients of the theory of consolidation,'' J. Appl. Mech. **24**, 1594–1601.

Bonnet, G. (**1987**). ''Basic singular solutions for a poroelastic medium in the dynamic range,'' J. Acoust. Soc. Am. **82**, 1758–1762.

Brebbia, C. A., Telles, J. C. F., and Wrobel, L. C. (**1984**). *Boundary Element Techniques* (Springer-Verlag, Berlin).

Cruse, T., and Rizzo, F. (**1968**). ''A direct formulation and numerical solution of the general transient elastodynamic problem,'' J. Math. Anal. Appl. **22**, 244–259.

Davies, B., and Martin, B. (**1979**). ''Numerical inversion of the Laplace transform: a survey and comparison of methods,'' J. Comput. Phys. **33**, 1–32.

Felippa, C. A. (**1980**). ''Top-down derivation of doubly asymptotic approximations for structure-fluid interaction analysis,'' *Innovative Numerical Analysis for the Engineering Sciences*, edited by R. P. Shaw *et al.* (University of Virginia, Charlottesville), pp. 79–88.

Geers, T. L. (**1971**). ''Residual potential and approximate methods for three-dimensional fluid-structure interaction problems,'' J. Acoust. Soc. Am. **49**, 1505–1510.

Geers, T. L. (**1978**). ''Doubly asymptotic approximations for transient motions of submerged structures,'' J. Acoust. Soc. Am. **64**, 1500–1508.

Geers, T. L., and Chen, J. (**1993**). ''On nonreflecting boundaries for poroelastic media,'' Proceedings IABEM 93, Braunschweig, Germany, edited by H. Antes (unpublished).

Geers, T. L., and Felippa, C. A. (**1983**). ''Doubly asymptotic approximations for vibration analysis of submerged structures,'' J. Acoust. Soc. Am. **73**, 1152–1159.

Geers, T. L., and Lewis, B. A. (**1997**). ''Doubly asymptotic approximations for transient elastodynamics,'' Int. J. Solids Struct. **34**, 1293–1305.

Geers, T. L., Oberai, A. A., and Yen, C.-L. (**1997**). ''Transient wave-exitation of a spherical shell embedded in an elastic medium,'' submitted for publication in J. Appl. Mech.

Geers, T. L., and Zhang, P. Z. (**1988**). ''Doubly asymptotic approximations for electromagnetic scattering problems,'' *Boundary Element Methods in Applied Mechanics*, edited by M. Tanaka and T. A. Cruse (Pergamon, Oxford), pp. 357–369.

Geers, T. L., and Zhang, P. Z. (**1994**). ''Doubly asymptotic approximations for submerged structures with internal fluid volumes: formulation,'' J. Appl. Mech. **61**, 893–899.

Johnson, D. L. (**1984**). ''Recent developments in the acoustic properties of porous media,'' in *Frontiers in Physical Acoustics*, edited by D. Sette (North-Holland, Amsterdam), pp. 255–290.

Junger, M. C., and Feit, D. (**1986**). *Sound, Structures, and their Interaction* (MIT, Cambridge, MA), 2nd ed.

Matthews, I. C., and Geers, T. L. (**1987**). ''A doubly asymptotic, non-reflecting boundary for ground-shock analysis,'' J. Appl. Mech. **54**, 489–497.

Nicolas-Vullierme, B. (**1991**). ''A contribution to doubly asymptotic approximations: an operator top-down approach,'' J. Vib. Acoust. **113**, 409–415.

Norris, A. N. (**1985**). ''Radiation from a point source and scattering theory in a fluid-saturated porous solid,'' J. Acoust. Soc. Am. **77**, 2012–2023.

Pierce, A. D. (**1989**). *Acoustics* (Acoustical Society of America, Woodbury, NY), 2nd ed.

Underwood, P. G., and Geers, T. L. (**1981**). ''Doubly symptotic boundary-element analysis of dynamic soil-structure interaction,'' Int. J. Solids Struct. **17**, 687–697.

Wolf, J. J., and Song, Ch. (**1995**). ''Doubly asymptotic multi-directional transmitting boundary for dynamic unbounded medium-structure-interaction analysis,'' Earthquake Eng. Struct. Dyn. **24**, 175–188.

Zimmerman, C., and Stern, M. (**1993**). ''Scattering of plane compressional waves by spherical inclusions in a poroelastic medium,'' J. Acoust. Soc. Am. **94**, 527–536.

Zimmerman, C., and Stern, M. (**1994**). ''Boundary element solution of 3-D wave scattering problems in a poroelastic medium,'' Engr. Anal. Boundary Elements **12**, 223–240.

Phenomenon of leaky free waves in the modal response of a uniform cylinder

G. Maidanik and K. J. Becker

David Taylor Model Basin (NSWC/CD), Bethesda, Maryland 20084-5000

The phenomenon of leaky free waves in the modal response of a uniform shell is manifest if (1) in the absence of fluid loading the locus of the free waves as a function of frequency crosses from the subsonic region to the supersonic region across the sonic locus and (2) in the presence of fluid loading the fluid surface impedance at and in the vicinity of the sonic locus is high and the introduction of this fluid loading does not, in itself, prevent the free waves of concern from reaching the sonic locus. In this consideration the sonic locus remains fixed relative to the properties of the shell and, therefore, changes in the fluid loading, on a given shell, are induced by changes in the fluid density only. The phenomenon of leaky free waves in the modal response of a uniform cylinder is investigated. It is found that two distinct examples of leaky free waves are manifested. The first is associated with the flexural free waves that reside in a frequency range that lies above the critical frequency with respect to the speed of sound in the fluid, and the second is associated with the curvature free waves that reside in a frequency range that lies below the ring frequency. The phenomenon of leaky free waves in these two examples is computed, displayed, and discussed. [S0001-4966(97)04809-1]

PACS numbers: 43.20.Jr, 43.20.Gp, 43.20.Tb, 43.40.Ey [JEG]

INTRODUCTION

In a recent report and a recent paper the authors dealt with the analysis of the normalized modal response $\bar{V}_n(k,\omega)$ of **hybrid** and **natural** fluid-loaded cylinders subjected to a normalized modal external drive $\bar{P}_{en}(k,\omega)$, where (k) is the axial wave number variable, (ω) is the frequency variable and (n) is the circumference mode index. [1,2] Section II of Ref. 2 is to be considered an integral part of this Introduction. In particular, Eqs. (1)–(15) in Sec. II of Ref. 2 now become Eqs. (1)–(15) in the present paper. In this section the investigation is limited to situations for which $(\omega_r/\omega_c)\ll 1$; to the lower mode indices, $n \lesssim 9$; and to the spectral range $\{0,(2\omega_r/\omega_c)\}\lesssim\{(ak),(\omega/\omega_c)\}\lesssim\{75,0.6\}$, where (a) is the radius of the cylinder, (ω_r) is the ring frequency, and (ω_c) is the critical frequency of the flexural free waves with respect to the speed of sound (c). [In this paper, as in Ref. 2, the ratio of the speed of sound (c) to the longitudinal speed (c_l), in the plating of the cylinder, remains fixed. Therefore, changes in the characteristic impedance of the fluid, on a cylinder with specific material properties, are induced by changes in the fluid density (ρ) only.] As much as this spectral range is of interest, there is a compelling investigative interest to extrapolate the range at both extremities. Of chief concern in this paper is the extrapolation that involves the phenomenon of ''leaky free waves.'' As can be inferred, here the phenomenon of leaky free waves is manifest and investigated in the normalized spectral domain; namely, either in the $\{(ak),(\omega/\omega_c)\}$-domain or in the $\{(ak),(\omega/\omega_r)\}$-domain. [3] Since this phenomenon with respect to the ''flexural free waves'' at the higher frequency range, where $(\omega/\omega_c)\gtrsim 1.0$, is reasonably familiar, whereas

this phenomena with respect to the ''membrane free waves'' at the lower frequency range, where $(\omega/\omega_r)\lesssim 2.0$, is hardly familiar, the story in this paper begins with extrapolation of the investigation into the higher frequency range just defined. [3–5] Under the conditions that $(\omega_r/\omega_c)\lesssim 1$ and $n \lesssim 9$, the phenomenon of leaky flexural free waves occurs in the partial response of a panel, and the modal responses of a hybrid cylinder and a natural cylinder. [1] Indeed, the literature deals with this phenomenon in all three of these shells. [3–5] Moreover, the differences in this phenomenon with a specific shell form are negligible; once it is accounted for in one shell, it is well nigh accounted for in any of the others. [6] On the other hand, it is recognized that neither the panel nor the hybrid cylinder exhibit membrane free waves and, therefore, only a natural cylinder is relevant to investigating the phenomenon of leaky free waves at the lower frequency range, where $(\omega/\omega_r)\lesssim 2.0$. [1,2] In addition, the present paper deals with uniform shells only; ribbed shells are dealt with in a companion paper that follows.

Accepting that ''familiarity breeds contempt,'' the leaky flexural free waves phenomenon at the higher frequency range, where $(\omega/\omega_c)\gtrsim 1.0$, is briefly repeated here so that the subsequent comparison with the leaky membrane free waves phenomenon at the lower frequency range, where $(\omega/\omega_r)\lesssim 2.0$, can be facilitated.

At this early stage a word or two about the figures presented in this paper may be useful. Typically, the figures depict the normalized quantity $|\bar{S}(k,\omega)|$ ''as a function of (ak) in a frequency waterfall format.'' [7] This quantity is displayed in terms of the quantity $\bar{S}'_q(ak)$ as a function of (ak). These quantities are related in the manner

$$\bar{S}'_q(ak) = \log\left[|\bar{S}\{k, \omega_0 + (q+1)\Delta\omega\}|10^q\right.$$
$$= q + \log|\bar{S}\{k, \omega_0 + (q+1)\Delta\omega\}|,$$

where (ω_0) is the frequency lower bound, $\omega_0 \lesssim \omega$; $(\Delta\omega)$ is an incremental frequency band; and (q) is an integer bounded by $0 \leq q \leq q_f$. As stated in the figure captions in this paper, the $\bar{S}(k,\omega)$, in $\bar{S}'_q(ak)$, is the normalized modal response $\bar{V}_{\infty n}(k,\omega)$. This equation states that the frequency is identified on the stepwise frequency axis by the frequency pseudonym 10^q, which locates the origin of the curve pertaining to the normalized frequency $(\omega/\omega_c) = [(\omega_0/\omega_c) + (q+1)\Delta(\omega/\omega_c)]$. It follows that the number of curves in a figure is equal to the ''upper bound integer'' (q_f) plus one and that (ω_c) is merely a convenient normalizing frequency. The spectral range for a display is defined by

$$\{(ak_0), (\omega_0/\omega_c)\} \lesssim \{(ak), (\omega/\omega_c)\} \lesssim \{(ak_f), (\omega_f/\omega_c)\},$$

yielding an estimate of q_f in the form $[(\omega_f/\omega_c) - (\omega_0/\omega_c)] \approx (q_f + 1)\Delta(\omega/\omega_c)$. For example, from Fig. 1(a) it is deduced that $60 < q_f < 70$; indeed, in this figure $(\omega_0/\omega_c) = 0.47$, $q_f = 65$, and $\Delta(\omega/\omega_c) = 0.03$, where the normalizing frequency (ω_c) is identified to be the critical frequency of the flexural free waves. Analogously, from Fig. 6(a) it is deduced that $98 < q_f < 103$; indeed, in this figure $(\omega_0/\omega_r) = 0$, $q_f = 101$, and $\Delta(\omega/\omega_r) = 0.014$, where the normalizing frequency (ω_r) is identified to be the ring frequency of the cylinder. Situations arise in which a feature in a display may be accentuated were the number of the displayed curves made more sparse; e.g., by successively displaying only one out of a fixed number of curves. This specific procedure is used in this paper; e.g., in Fig. 1(a) every other curve is omitted and, therefore, only 33 out of 66 curves are shown. Analogously, in Fig. 6(a) two curves out of three are omitted, and therefore, only 34 out of 102 curves are shown. Other variations on the display theme may be similarly defined and introduced when one or another feature calls. The sonic locus positions in the figures are determined, in the context of the paper, in the compatible form

$$(ak)^2 = [(\omega_0/\omega_c) + (q+1)\Delta(\omega/\omega_c)]^2(ak_c)^2 - n^2],$$
$$k_c = (\omega_c/c),$$

where, again, (ω_c) is merely a convenient normalizing frequency and (q) is the integer just defined with respect to the frequency waterfall format.

I. PHENOMENON OF LEAKY FLEXURAL FREE WAVES IN A FLUID-LOADED UNIFORM CYLINDER

From Eq. (13a) the absolute value of the normalized modal response $|\bar{V}_{\infty n}(k,\omega)|[\equiv|\bar{G}_{\infty n}(k,\omega)|]$ is evaluated as a function of (ak) in a frequency waterfall format.[1,7] The evaluation in this section is a ''zoom on'' the higher spectral range in which the critical frequency is approached and surpassed. This spectral range starts at the upper limits of the

range previously considered in Ref. 2 and extends the normalized frequency range to an octave above the critical frequency; the zoom on in the higher spectral range is thus defined: $\{70, 0.5\} \lesssim \{(ak), (\omega/\omega_c)\} \lesssim \{150, 2.5\}$. [As already mentioned, in this range the partial (or modal) response is substantially invariant to the form of the shell.] The phenomenon of leakage of free waves is advantageously introduced via a situation in which it is absent; a situation of light fluid loading. Note that a moderate fluid loading is characterized by the standard fluid-loading parameter $\varepsilon_c = 10^{-2}$, a light fluid loading is characterized by $\varepsilon_c \lesssim 10^{-4}$, and heavy fluid-loading parameter by $\varepsilon_c \gtrsim 10^{-1}$. Also note that the change in the fluid-loading parameter, in the context of this paper, is effected by the fluid density (ρ); the fluid speed of sound (c) remains unchanged. Figure 1(a)–(c) depicts $\bar{V}_{\infty n}(k,\omega)$, as a function of (ak) in a frequency waterfall format, in the higher spectral range just defined, for three values of the mode index; $n = 0, 1$, and 9, respectively, and for light fluid loading; $\varepsilon_c = 10^{-4}$. The sonic locus is superimposed on these figures so that its location can be identified. The flexural locus, constituted by the ridges and associated peaks, defines the flexural free waves in these figures. Clearly, as the normalized frequency increases from $(\omega/\omega_c) = 0.5$, the flexural locus converges on the sonic locus from the subsonic region. When (ω/ω_c) reaches unity, this convergence is completed. As (ω/ω_c) is increased beyond unity, the flexural locus crosses the sonic locus unimpeded and begins to diverge from it into the supersonic region. The flexural locus diverges more and more from the sonic locus as (ω/ω_c) further increases on to 2 and beyond.

Figure 2(a)–(c) repeats Fig. 1(a)–(c), respectively, except that the standard fluid-loading parameter is restored; $\varepsilon_c = 10^{-2}$. The sonic locus, defining a spectral region of high fluid surface impedance, is discernible in these figures. Even for this moderate fluid loading, the fluid surface impedance is high enough to imprint the sonic locus on these figures. This high fluid surface impedance at the sonic locus establishes a barrier which impedes the crossing of the flexural free waves from the subsonic region into the supersonic region. As the normalized frequency (ω/ω_c) increases past unity, the flexural locus is held at bay adjacent to the barrier on the subsonic side. As the normalized frequency is further increased, however, a gradual commencement of leakage of flexural free waves through the barrier at the sonic locus into their rightful place in the supersonic region occurs. The leakage is substantially completed when (ω/ω_c) is 2.

The higher fluid loading, which is essential to the manifestation of the phenomenon of leaky flexural free waves, introduces an additional phenomenon. Clearly, the flexural free waves that are in the supersonic region are efficient radiators compared with those in the subsonic region. When fluid loading is light, this difference is not significant since radiation damping, even for the supersonic-flexural free waves, is not high enough to compete with the inherent mechanical damping. However, when fluid loading is moderate, the radiation damping of the supersonic flexural free waves becomes significant, as can be verified by comparing Figs. 1 and 2.

Figure 3(a1), (b), and (c) repeats Fig. 2(a)–(c), respec-

Zoom Ginbar, N=0, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-4

(a)

Zoom Ginbar, N=1, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-4

(b)

Zoom Ginbar, N=9, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-4

(c)
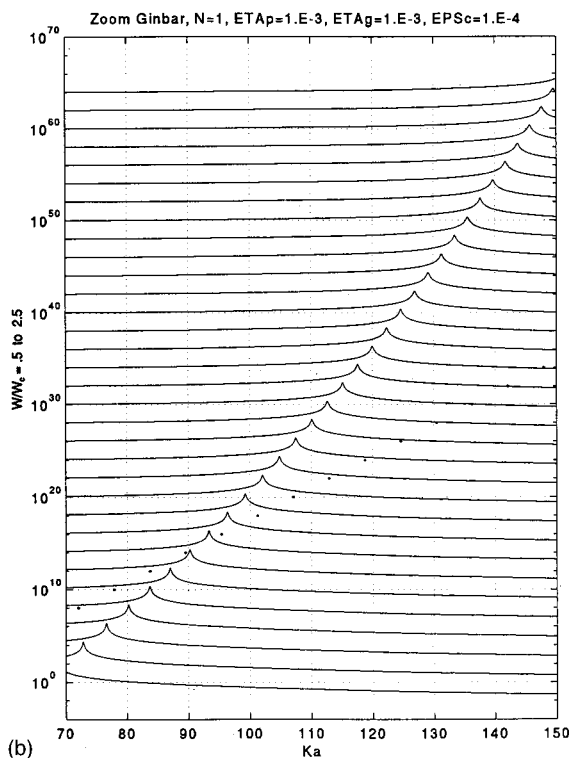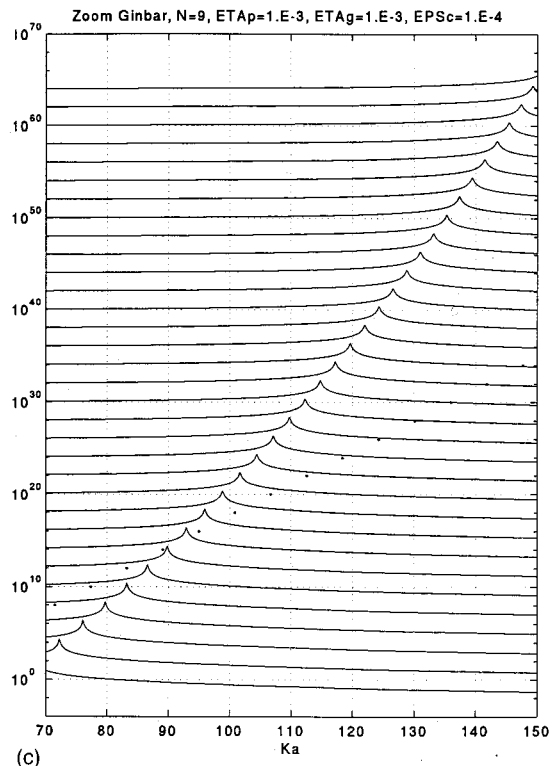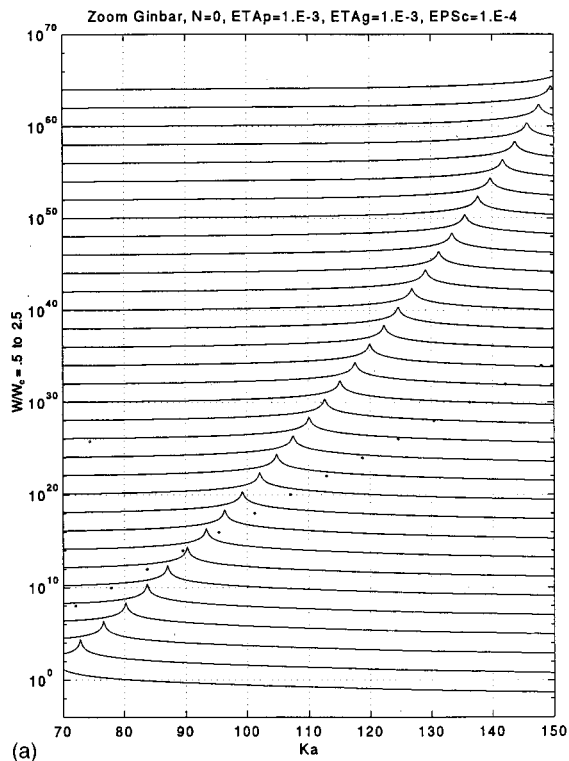
FIG. 1. The normalized modal response $\overline{V}_{\infty n}(k,\omega)$ as a function of (ak) in a frequency waterfall format pertaining to the higher spectral range. Standard parametric values are used, except that the fluid loading parameter (ε_c) is changed from the standard value of 10^{-2} to 10^{-4} and the mode index (n) is as specified. Superimposed is the sonic locus defined by $(ak)^2 = [(\omega/\omega_c)^2 (ak_c)^2 - n^2]$; $k_c = (\omega_c/c)$ and is represented by discrete dots. (a) $n=0$. (b) $n=1$. (c) $n=9$.

tively, except that the fluid-loading parameter is changed from the moderate standard value of 10^{-2} to the high value of 10^{-1}. Comparing Figs. 3(a1), (b), and (c) and 2(a)–(c), respectively, shows the increased effectiveness of the fluid-loading barrier in the former set of figures and the accompanied decrease in the leakage of flexural free waves across this

barrier. The more gradual leakage across the more effective barrier results in a wider frequency band to complete the leakage. Thus, this frequency band in Fig. 3(a1), (b) and (c) is wider than in Fig. 2(a)–(c), respectively. Indeed, the completion of the leakage in Fig. 3(a1), (b), and (c) takes place beyond the frequency range used for these figures, as

Zoom Ginbar, N=0, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-2

(a)

Zoom Ginbar, N=9, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-2

(c)

Zoom Ginbar, N=1, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-2

(b)
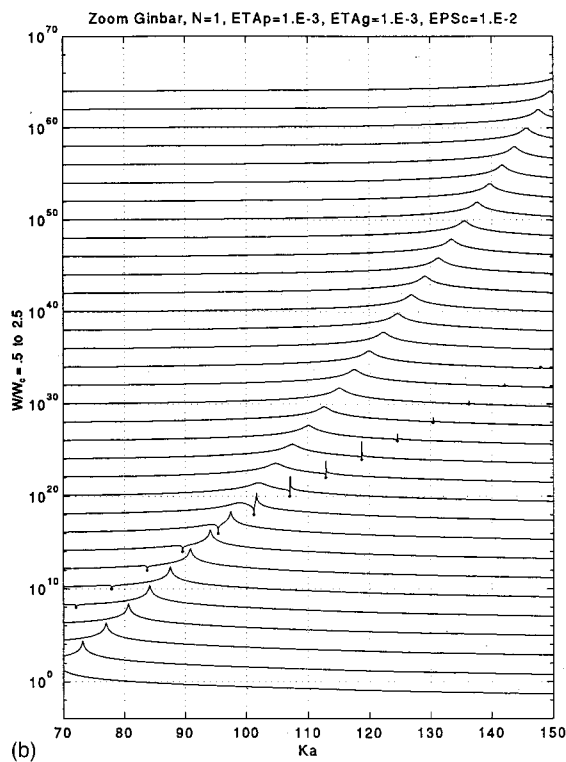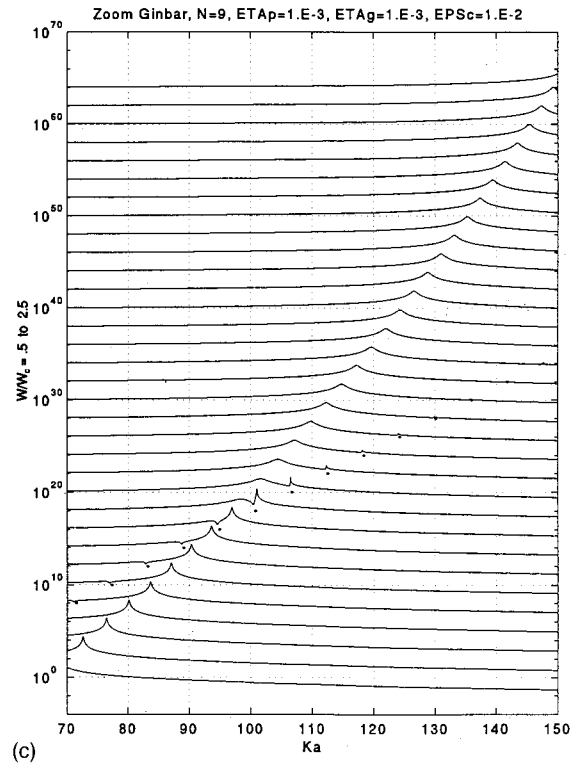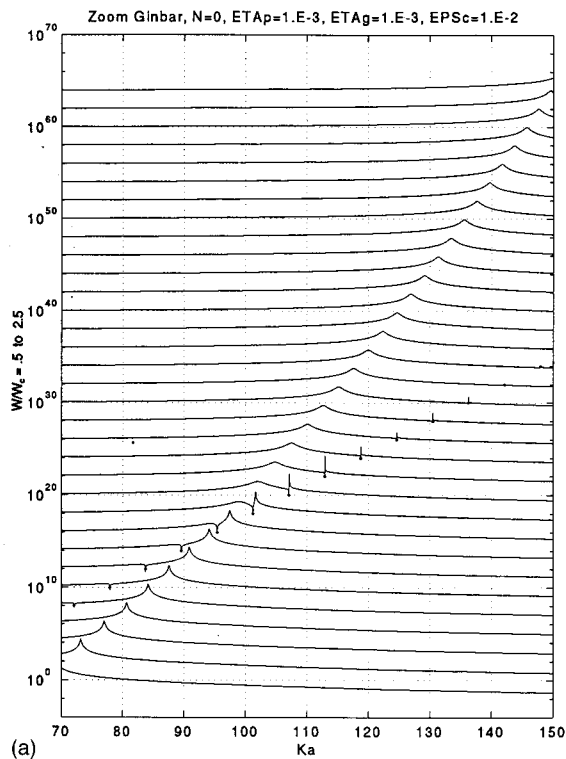
FIG. 2. As in Fig. 1 except that the fluid-loading parameter (ε_c) is restored to the standard value of 10^{-2}. (a) $n=0$. (b) $n=1$. (c) $n=9$.

Fig. 3(a2) attests. Moreover, as expected, the radiation damping of the flexural free waves, in the supersonic region, is higher in Fig. 3(a1), (b), and (c) than in Fig. 2(a)–(c), respectively.

The phenomenon of leaky free waves is thus predicated on the existence of free waves that, in the absence of fluid loading, transit, as a function of frequency, across the sonic locus. In addition, the fluid loading needs to be high enough to establish an effective barrier that substantially impedes this crossing. The flexural free waves in a panel, in a hybrid cylinder, and in a natural cylinder satisfy the first of these requirements. Are there other types of free waves on a shell
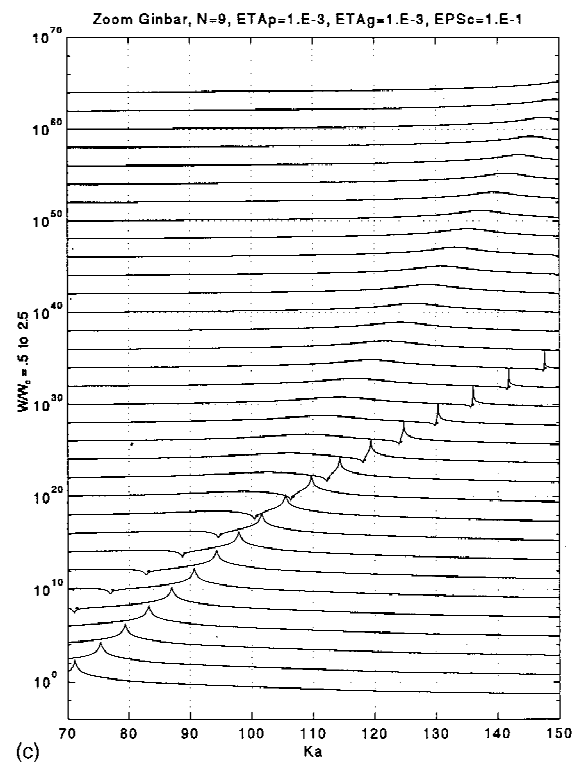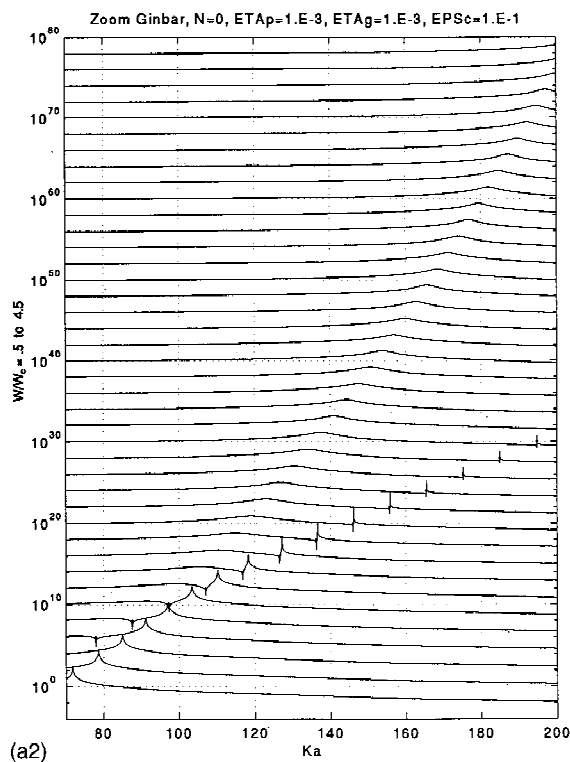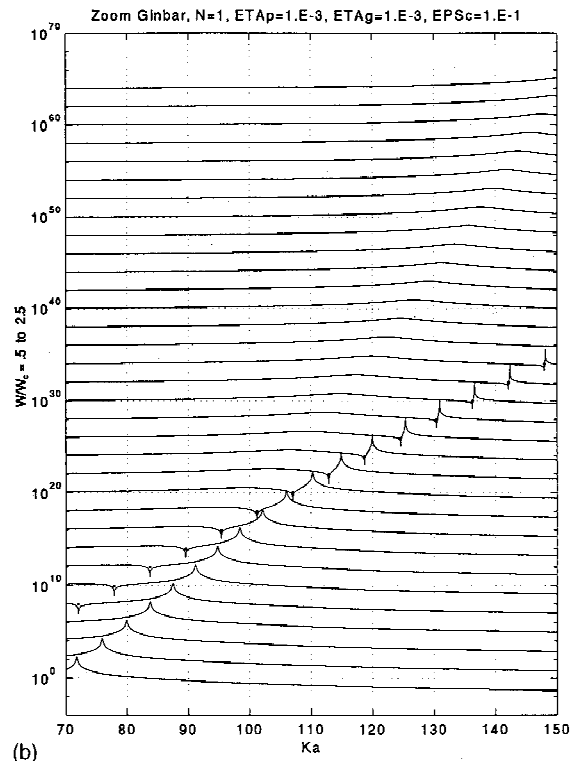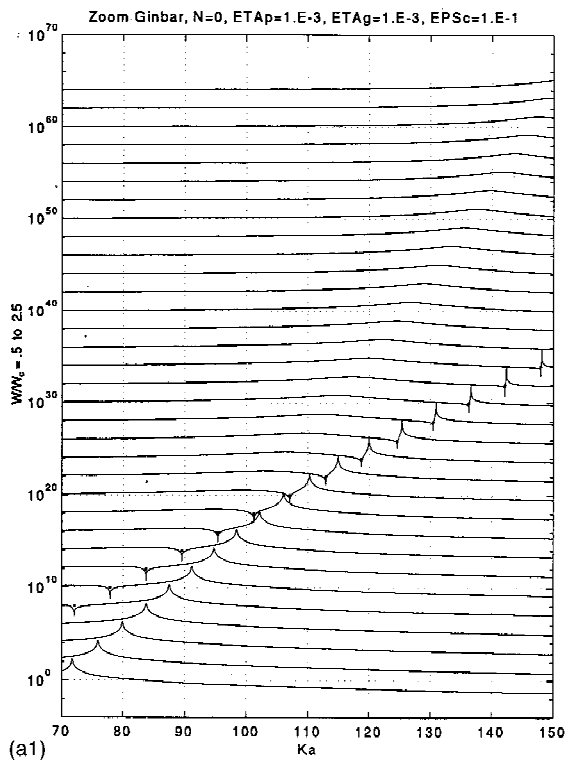
FIG. 3. As in Fig. 1 except that the fluid-loading parameter (ε_c) is changed from the value of 10^{-4} to 10^{-1}. (a) $n=0$. (a2) As in (a1) except that the spectral range is extended. (b) $n=1$. (c) $n=9$.

that similarly satisfy this first requirement? The natural cylinder is the only shell of these mentioned that needs to be examined in light of this question; the other two shells support, by definition, flexural free waves only, whereas the natural cylinder supports, in addition, membrane free waves.

II. BEHAVIOR IN THE LOWER FREQUENCY RANGE OF THE MEMBRANE FREE WAVES IN A NATURAL UNIFORM CYLINDER

Again, advantageously the phenomenon of leakage in the membrane free waves is introduced via a situation in
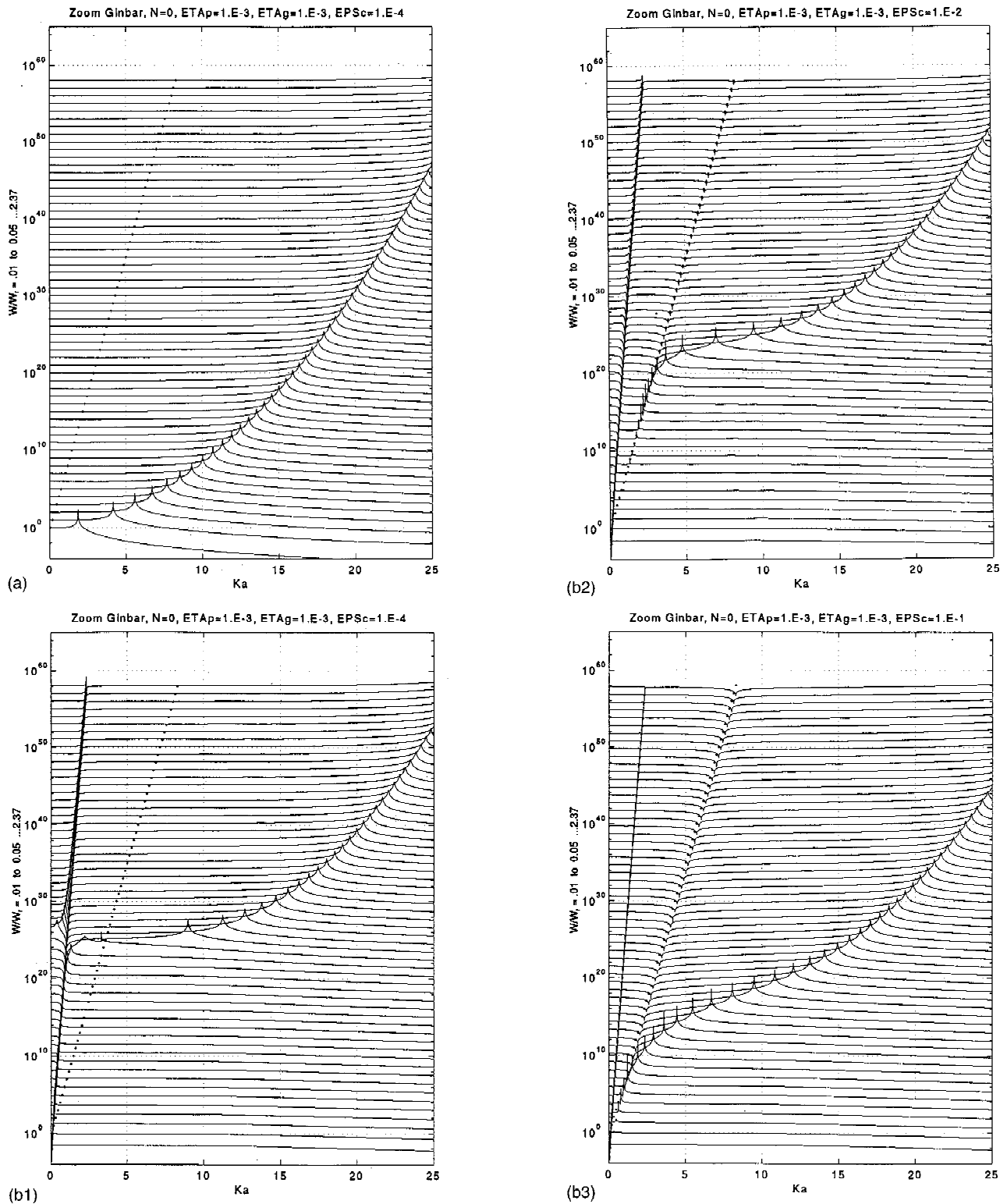
FIG. 4. The normalized modal response $\bar{V}_{\infty n}(k,\omega)$ as a function of (ak) in a frequency waterfall format pertaining to the lower spectral range. Standard parametric values are used, except that the mode index (n) is zero and the fluid-loading parameter (ε_c) is as specified. Superimposed is the sonic locus represented by discrete dots. (a) A hybrid cylinder; $\varepsilon_c = 10^{-4}$. (b1) A natural cylinder; $\varepsilon_c = 10^{-4}$. (b2) $\varepsilon_c = 10^{-2}$. (b3) $\varepsilon_c = 10^{-1}$.

which it is absent: a situation of light fluid loading. From Eq. (13a) the absolute value of the normalized modal response $|\bar{V}_{\infty n}(k,\omega)| = [\equiv |\bar{G}_{\infty n}(k,\omega)|]$ is evaluated as a function of (ak) in a frequency waterfall format.[1,7] This evaluation is confined to the lower spectral range defined by $\{0,0\}$

$\lesssim \{(ak),(\omega/\omega_r)\} \lesssim \{25,2.4\}$. The modal response, pertaining to the membrane free wave, is more characteristically identified in a frequency waterfall representation that employs a frequency normalization by the ring frequency (ω_r), rather than by the critical frequency (ω_c). Figures 4(a) and (b1)
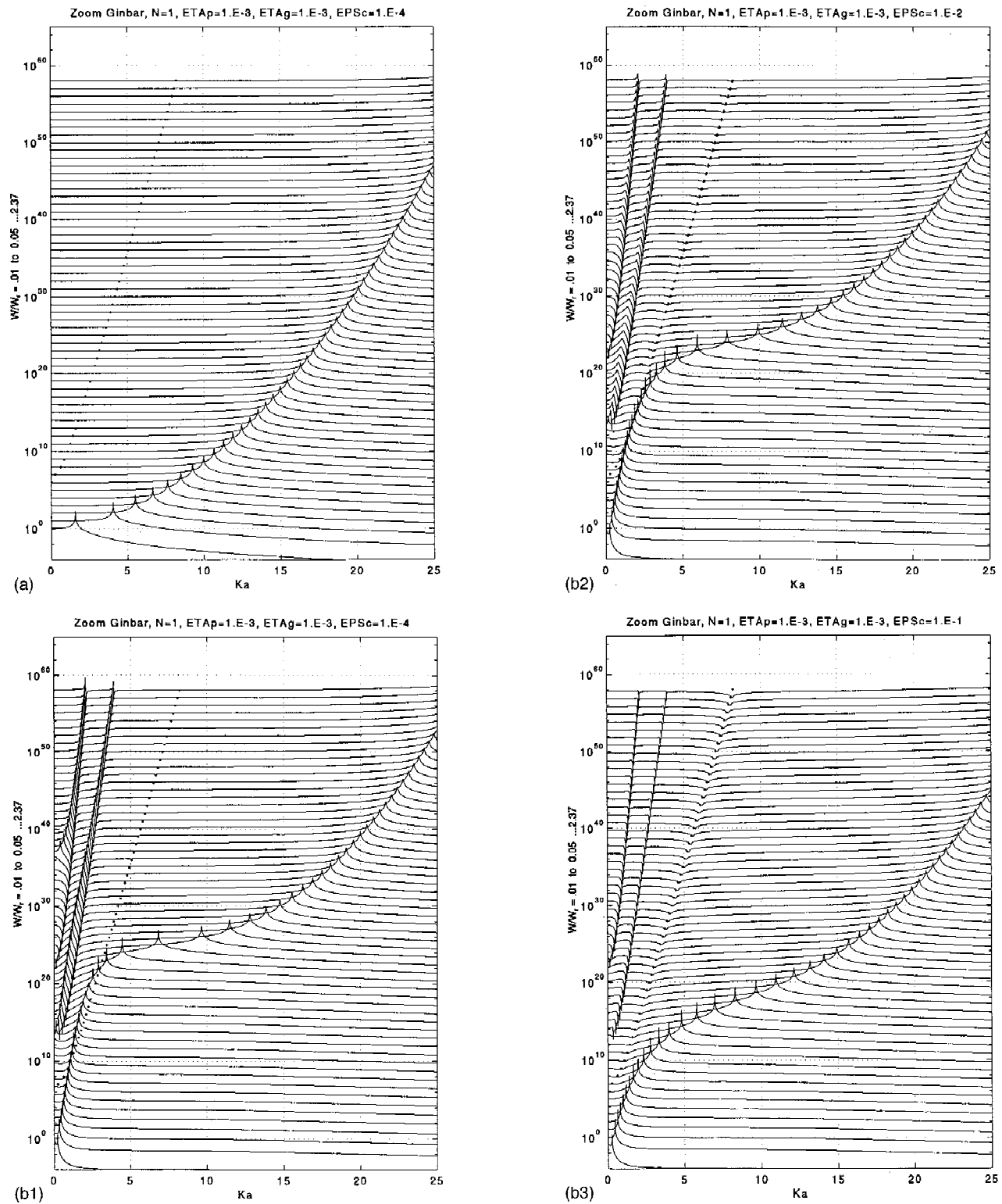
FIG. 5. As in Fig. 4 except that the mode index (n) is unity. (a) A hybrid cylinder; $\varepsilon_c = 10^{-4}$. (b1) A natural cylinder; $\varepsilon_c = 10^{-4}$. (b2) $\varepsilon_c = 10^{-2}$. (b3) $\varepsilon_c = 10^{-1}$.

and 5(a) and (b1) are evaluated under standard parametric values except that the fluid-loading parameter (ε_c) is changed from the standard value of 10^{-2} to 10^{-4}, and in the first set of figures the mode index(n) is zero and in the second and standard value of unity prevails. Figures 4(a) and 5(a) pertain to a hybrid cylinder and the corresponding Figs.

4(b1) and 5(b1) pertain to a natural cylinder. Comparing Figs. 4(a) and 5(a) with Figs. 4(b1) and 5(b1), respectively, shows that the membrane free waves are composed of longitudinal and curvature free waves in Fig. 4(b1) and of longitudinal, shear, and curvature free waves in Fig. 5(b1).

Figure 4(b1), for which the mode index (n) is equal to

zero, exhibits the birth of the longitudinal free waves at $\{(ak),(\omega/\omega_r)\}\simeq\{0,1.0\}$. In the normalized frequency range below unity, $(\omega/\omega_r)\lesssim 1$, the longitudinal locus is absent. Close to its spectral place of birth, the speed of the longitudinal free waves is high, decreasing rapidly, and converging onto the longitudinal speed (c_l) as the normalized frequency (ω/ω_c) increases past unity; the longitudinal speed (c_l) is independent of frequency. On the outside and adjacent to the longitudinal locus lies an ''antilongitudinal locus'' that is defined by valleys and associated nadirs. This antilongitudinal locus maintains the longitudinal speed and it continues uninterrupted as the normalized frequency (ω/ω_r) is decreased below unity on to the spectral origin at $\{(ak),(\omega/\omega_r)\}=\{0,0\}$. Again, there are no shear and anti-shear loci patterns for $n=0$.[1] Finally, the flexural free waves at the lower frequency range, where $(\omega/\omega_r)\lesssim 2.0$, are superseded by curvature free waves. The speed of the curvature free waves surpasses that of the corresponding flexural free waves more and more as (ω/ω_r) decreases below 2 and on to unity [cf. Fig. 4(a)]. As (ω/ω_r) decreases further below unity, the speed of the curvature locus, which defines the curvature free waves, assumes the longitudinal speed, positioning itself on the outside of and adjacent to the antilongitudinal locus. The curvature free waves in that position are guided by the antilongitudinal locus all the way to the spectral origin. Since the longitudinal speed (c_l) is higher than the speed of sound (c), as defined in Eq. (15), the curvature free waves emerge as a potential candidate to exhibit the phenomenon of leaky free waves. Confluently, this description designates (ω_r) to be the ''critical frequency,'' with respect to the speed of sound (c), of the curvature free waves pertaining to the zeroth-mode index.[2]

Figure 5(b1), for which the mode index (n) is equal to the standard value of unity, exhibits both the longitudinal locus and the shear locus.[1] Moreover, each is accompanied on the outside by an adjacent antilocus; an antilongitudinal locus and an antishear locus. In Fig. 5(b1) neither of these loci nor the accompanying antiloci extend beyond $(\omega/\omega_r)\simeq\frac{1}{2}$; the birth of the longitudinal locus is at a higher normalized frequency than in Fig. 4(b1) and the birth of the shear locus in Fig. 5(b1) is at $(\omega/\omega_r)\simeq\frac{1}{2}$. The longitudinal and antilongitudinal loci asymptotically approach the longitudinal speed (c_l) and the shear and anti-shear loci asymptotically approach the shear speed (c_s); none of these loci cross or even closely approach the sonic locus and, therefore, none is a potential candidate to generate leaky free waves. On the other hand, again at the lower frequency range, where $(\omega/\omega_r)\lesssim 2.0$, the flexural free waves are superseded by curvature free waves. The speed of the curvature free waves surpasses that of the corresponding flexural free waves more and more as (ω/ω_r) decreases below 2 and on to unity [cf. Fig. 5(a)]. As (ω/ω_r) decreases further below unity, the speed of the curvature locus become more and more independent of frequency as it makes its way toward the spectral origin at $\{(ak),(\omega/\omega_r)\}\cong\{0,0\}$. In Fig. 5(b1) this asymptotic speed is approximately the shear speed (c_s), ha!, may be a shade lower. In addition, the off-set in the sonic locus when $n=1$ helps render the curvature free waves in

Fig. 5(b1) only marginally a potential candidate to exhibit the phenomenon of leaky free waves.

Clearly, for a mode index (n) that exceeds unity, the curvature of the cylinder is even weaker than it is for a mode index of unity and, therefore, the conversion of the flexural free waves into curvature free waves is less effective. Under the parametric values specified in Eq. (15), which are central in the evaluations in this paper, the phenomenon of leaky waves is not expected for mode indices that exceed unity, is only marginally expected for a mode index of unity, and is definitely expected for a mode index of zero.

III. PHENOMENON OF LEAKY CURVATURE FREE WAVES IN A FLUID-LOADED NATURAL UNIFORM CYLINDER

A ''zoom on region'' where the phenomenon of leaky curvature free waves can be effectively and conveniently observed is, in the lower spectral range, defined by $\{0,0\}\lesssim\{(ak),(\omega/\omega_r)\lesssim\{4.0,1.4\}$; this spectral range is sufficient to expose the essential features in the phenomenon. Figure 6(a) is the zoom on region taken out of Fig. 4(b1); in these figures the mode index (n) is zero, and the fluid-loading parameter (ε_c) is light at 10^{-4}. In Fig. 6(a), as in Fig. 4(b1), the leaky free waves phenomenon is absent. Figure 6(b) repeats Fig. 6(a) except that the standard fluid-loading parameter is restored; $\varepsilon_c=10^{-2}$. Observe that in general the longitudinal free waves are substantially subdued by radiation damping that the fluid loading encourages; after all, the longitudinal locus lies in the supersonic spectra region. Moreover, the higher fluid loading in Fig. 6(b), as compared with that in Fig. 6(a), tends to delay and mollify the transition of curvature free waves from the flexural free waves that they supersede (cf. the Appendix). As the normalized frequency (ω/ω_r) is decreased past unity, the curvature free waves are held at bay by the fluid-loading barrier at the sonic locus. Further decrease in (ω/ω_r) introduces a gradual leakage of curvature free waves that gradually reestablishes the curvature locus at its rightful position adjacent and on the outside of the antilongitudinal locus, [cf. Fig. 6(a)]. The leakage is complete when the normalized frequency (ω/ω_r) reaches the value of $\frac{1}{3}$ [cf. Sec. I and, in particular, the comparison between Figs. 1(a) and 2(a)]. Figure 6(c) repeats Fig. 6(b) except that the standard fluid parameter is changed to the higher value of 10^{-1}. Compared with Fig. 6(b), the features in Fig. 6(c) that are associated with the influences of fluid loading and, in particular, with the phenomenon of leaky curvature free waves are more pronounced. Again, the normalized frequency band to complete the leakage is wider and extends further toward the spectral origin at $\{(ak),(\omega/\omega_r)\}=\{0,0\}$ (compare Figs. 2 and 3).

A zoom on region where the marginal phenomenon of leaky curvature free waves can be effectively and conveniently observed is in the same spectral range employed in Fig. 6. Figure 7(a) depicts the zoom on region taken out of Fig. 5(b1); in these figures the mode index (n) is unity and the fluid-loading parameter (ε_c) is light at 10^{-4}. In Fig. 7(a), as in Fig. 5(a), the leaky free wave phenomenon is absent even in a marginal form. Figure 7(b) repeats Fig. 7(a) except that the fluidloading parameter is restored to the standard

Zoom Ginbar, N=0, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-4

(a)

Zoom Ginbar, N=0, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-1

(c)

Zoom Ginbar, N=0, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-2

(b)
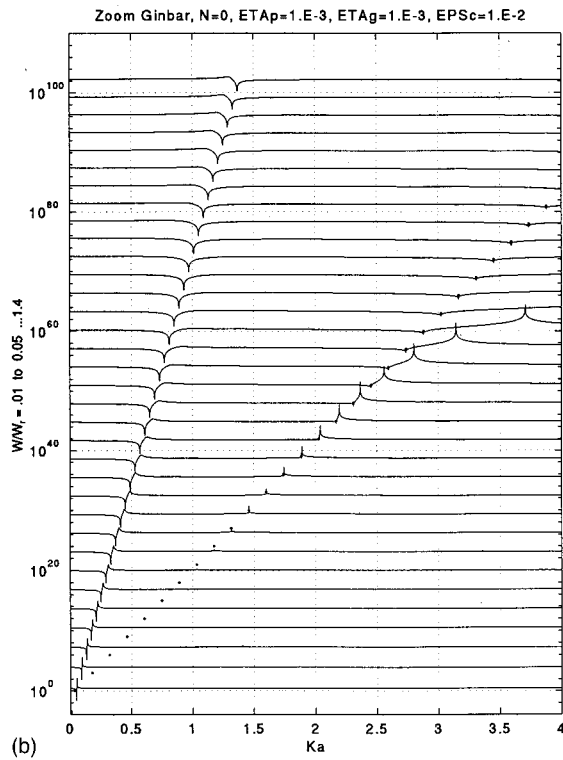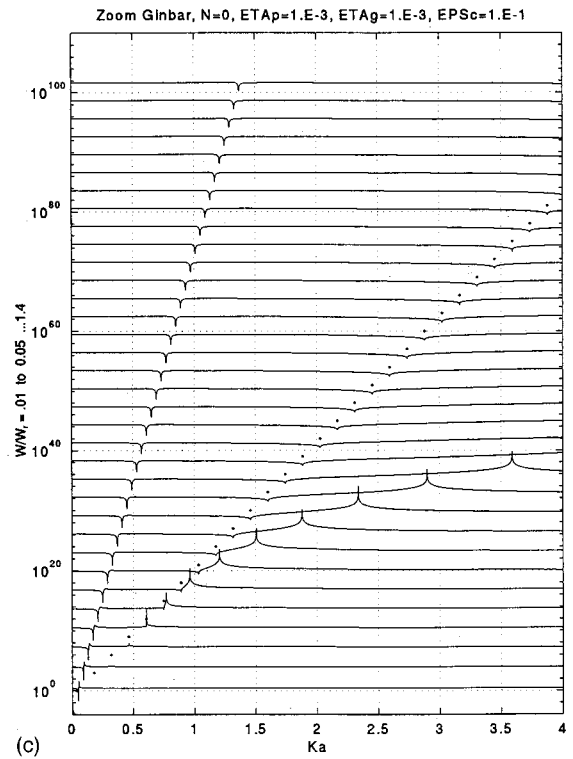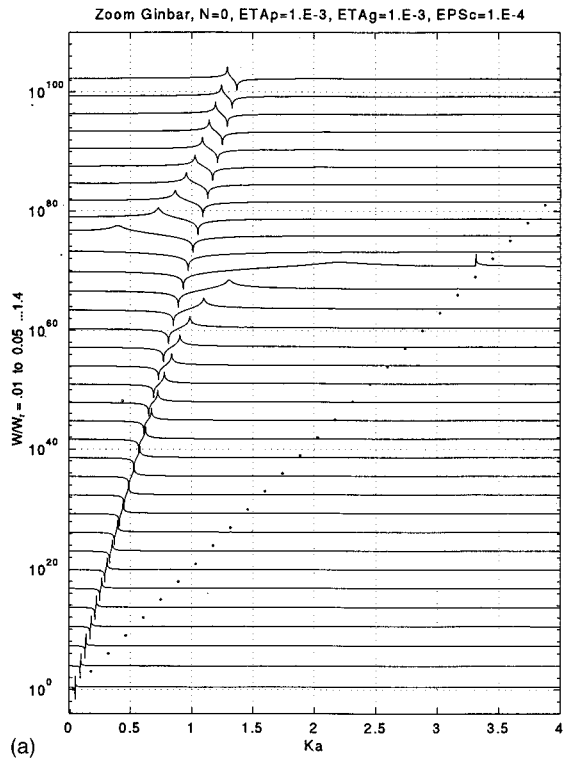
FIG. 6. The normalized modal response $\overline{V}_{\infty n}(,\omega)$ as a function of (ak) in a frequency waterfall format pertaining to the zoom on region in the lower spectral range. Standard parametric values are used, except that the mode index (n) is zero and the fluid loading parameter (ε_c) is as specified. Superimposed is the sonic locus represented by discrete dots. (a) $\varepsilon_c = 10^{-4}$. (b) $\varepsilon_c = 10^{-2}$. (c) $\varepsilon_c = 10^{-1}$.

value of 10^{-2}. Observe that, in general, the longitudinal and the shear free waves are substantially subdued by radiation damping that is caused by the increase in fluid loading; Eq. (15) dictates that both the longitudinal and shear loci lie in the supersonic spectral region; $c_l > c_s > c$. Moreover, the higher fluid loading and mode index in Fig. 7(b) as compared with Figs. 7(a) and 6(b), respectively, conspire to delay and mollify the transition of the curvature free waves from the

flexural free waves that they supersede. As (ω/ω_r) is decreased past unity, the curvature free waves are held at bay by the fluid-loading barrier at the sonic locus [cf. Fig. 7(a)]. There is no true leakage; the curvature locus is merely held adjacent and on the subsonic side of the sonic locus all the way to the spectral origin. Figure 7(c) repeats Fig. 7(b) except that the standard fluid-loading parameter is changed to the higher value of 10^{-1}. Again, compared with Fig. 7(b),

Zoom Ginbar, N=1, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-4

(a)

Zoom Ginbar, N=1, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-2

(b)

Zoom Ginbar, N=1, ETAp=1.E-3, ETAg=1.E-3, EPSc=1.E-1

(c)
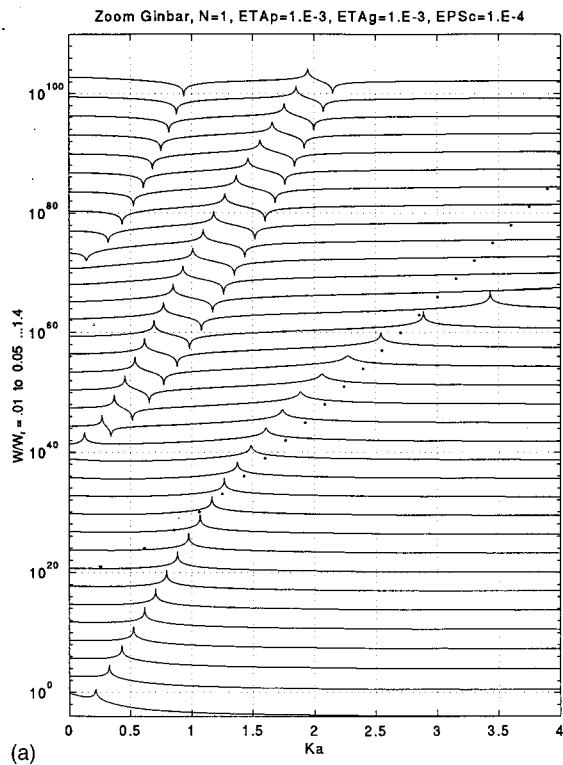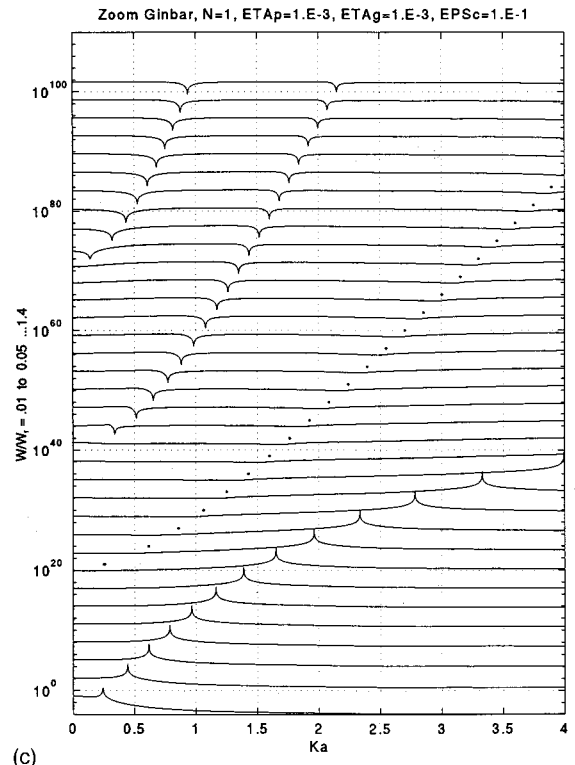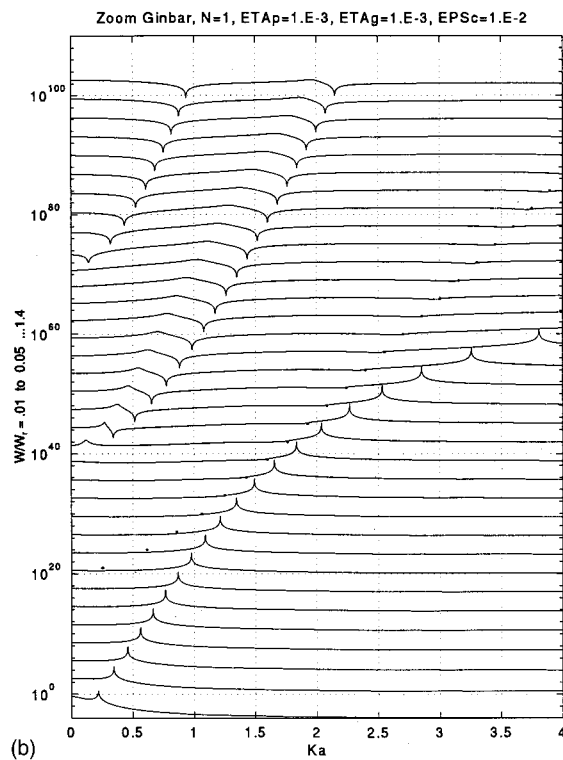
FIG. 7. As in Fig. 6 except that the mode index (n) is unity. (a) $\varepsilon_c = 10^{-4}$. (b) $\varepsilon_c = 10^{-2}$. (c) $\varepsilon_c = 10^{-1}$.

the features in Fig. 7(c) that are associated with the influences of fluid loading and, in particular, with the marginal phenomenon of leaky curvature free waves, are more pronounced. Indeed, the transition of the curvature free waves, from the flexural free waves that they supersede, is sufficiently delayed and mollified by the increased fluid loading that the phenomenon of leaky curvature free waves, in Fig. 7(c), is not even marginally manifested [cf. Fig. 6(b) and (c), and the Appendix].

APPENDIX

The mollifying effects of fluid loading on scattering and other structural response phenomena are well known.[1,4,5] It may then be helpful to briefly consider the role that fluid loading may play in the transition, as such, of flexural free waves into curvature free waves. Since this transition is more pronounced the lower the mode index is, only $n=0$ and $n=1$ are selected for the investigation in this Appendix.

Figure 4(b2) and (b3) repeats Fig. 4(b1) except that the fluid-loading parameters (ε_c) is changed from 10^{-4} to the standard value of 10^{-2} and to the higher value of 10^{-1}, respectively. Disregarding the phenomenon of leaky curvature free waves that is discussed in the text, these three figures clearly show that increases in fluid loading tend to delay and mollify the transition to curvature free waves from the flexural free waves that they supersede.[1,2] These effects are only slight from Fig. 4(b1) to Fig. 4(b2), but are substantial from Fig. 4(b1) to Fig. 4(b3). Nonetheless, even at the high value of 10^{-1} for the fluid parameter (ε_c), the delay and mollifying effects of fluid loading are not sufficient to scuttle the phenomenon of leaky curvature free waves, as Fig. 6(c) verifies.

Figure 5(b2) and (b3) repeats Fig. 5(b1) except that the fluid-loading parameter (ε_c) is changed from 10^{-4} to 10^{-2} and 10^{-1}, respectively. Again, disregarding the phenomenon of leaky curvature free waves that is discussed in the text, these three figures clearly show that increases in fluid loading tend to delay and mollify the transition to curvature free waves from the flexural free waves that they supersede.[1,2] Although these effects are only slight from Fig. 5(b1) to Fig. 5(b2), they are sufficient to render the curvature free waves subsonic, notwithstanding that at and in the vicinity of $(\omega/\omega_r) \simeq (\frac{1}{2})$ the curvature free waves are only just subsonic [cf. Fig. 7(b)]. On the other hand, the effects are substantial from Fig. 5(b1) to Fig. 5(b3). Indeed, the high fluid loading, in this case, renders the curvature free waves well nigh subsonic, as Fig. 7(c) verifies.

[1] G. Maidanik and K. J. Becker, ''Phenomena of aliasing and pass and stop bands in the drive in lieu of ribs on cylindrical shells,'' NSWCCD Report No. SIG-96/072-7030 (1996).

[2] G. Maidanik and K. J. Becker, ''Computation of the modal response of regularly ribbed cylinders,'' J. Acoust. Soc. Am. **101**, 2452–2472 (1997).

[3] 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).

[4] D. G. Crighton, ''The green function of an infinite, fluid loaded membrane,'' J. Sound Vib. **86**, 411–433 (1983); ''Transmission of energy down periodically ribbed elastic structures under fluid-loading,'' Proc. R. Soc. London, Ser. A **394**, 405–436 (1984).

[5] D. M. Photiadis, ''The propagation of axisymmetric waves on a fluid-loaded cylindrical shell,'' J. Acoust. Soc. Am. **88**, 239–250 (1990); J. F. M. Scott, ''The free modes of vibration of an infinite fluid-loaded cylindrical shell,'' J. Sound Vib. **125**, 241 (1988).

[6] F. Fahy, *Sound and Structural Vibration (Radiation, Transmission and Response)* (Academic, New York, 1985).

[7] G. Maidanik and J. Dickey, ''Velocity distributions on unloaded finitely and regularly ribbed membranes,'' J. Acoust. Soc. Am. **149**, 43–70 (1991); ''Response of regularly ribbed fluid loaded panels,'' J. Acoust. Soc. Am. **155**, 481–495 (1992).

On the overlapping acoustic resonances of a fluid-filled cavity: Schlieren visualization of an insonified circular-cylindrical shell

Paul A. Chinnery and Victor F. Humphrey
School of Physics, University of Bath, Bath BA2 7AY, United Kingdom

A schlieren technique is used to visualize the fluid cavity resonances of an insonified fluid-filled (brass) cylindrical shell in the high-frequency overlapping resonance regime, $kb > 30$, where b is the inner radius of the shell. Hybridization of modes occurs at frequencies where resonances interfere and the resulting dissymetrization of the wave fields are evident in the experimental images. Similar behavior is seen in theoretical predictions obtained using the normal mode series solution for a shell excited by a plane wave. At very high frequencies ($kb > 100$) the field in the cavity has characteristics that can be associated with a ray description of acoustic propagation; in these cases caustics are observed in the acoustic field. © *1997 Acoustical Society of America.*
[S0001-4966(97)06008-6]

INTRODUCTION

Resonances play an important part in the scattering of acoustic waves by discrete submerged elastoacoustic objects. Using the resonance scattering theory (RST)[1] the scattered field can be divided into two components; the resonances of the scattering target, and the nonresonant background. In the form function of a target the resonances manifest themselves as sharp changes superimposed upon the more smoothly varying background. Some information regarding the resonances can be obtained from the scattered field; either by studying the scattered time signal, or by examining the spectrum of the scattered field. However, if the target is a cylindrical shell and the inner cavity contains a liquid, the acoustic field inside the cavity can provide valuable information about the target resonances, many of which are associated directly with the fluid cavity itself. For fluid-filled cylindrical shells three classes of resonance have been identified;[2] resonances of the fluid column, resonances of the elastic shell, and Stoneley wave resonances. When insonified with a continuous-wave signal, resonances of each kind manifest themselves as standing waves within the cavity. Many of the vibrational modes of cylindrical structures could be identified if the acoustic field in the fluid column could be measured. Unlike the external (scattered) field, the field within the fluid cavity has received relatively little attention in the literature.

One method of visualizing acoustic fields in fluids is the schlieren technique.[3–7] This approach can be used to study the resonance and scattering behavior of cylindrical objects in detail.[8–13] The advantage of this technique is that it is noninvasive, allowing the field inside the fluid column to be monitored without disturbing it. The distinctive standing wave patterns within the fluid column are clearly seen with schlieren, allowing easy identification of the modes.

A low-frequency schlieren system permitting the visualization of pulsed and continuous acoustic fields down to about 100 kHz has been developed for studying scattering at low frequencies by discrete objects of reasonable size.[9] This system has been used to study the resonance behavior of circular cylindrical objects,[10] and elliptical-[11] and stadium-[13] shaped cavities. The location, identification, and imaging of resonance modes is particularly simple using this technique; the experimental system and interpretation of the images is described in greater detail elsewhere.[8]

The present paper discusses the application of this technique to the resonances of an insonified circular cylindrical shell at high frequencies, where resonances overlap and hybridization of modes occurs.

I. DESCRIPTION OF PROBLEM AND THEORY

For a thick metallic fluid-filled shell insonified by an acoustic wave (Fig. 1) the most abundant resonances are those associated with the fluid column. These resonances are the subject of this paper. For a lossless cavity the fluid column resonances are simply the eigenvibrations of the fluid column (with Neuman boundary conditions) and the resonance frequencies constitute the spectrum of eigenvalues of the cavity, each "resonance" having zero width. The eigenvibrations (normal modes) of such a circular cavity of radius b with rigid boundary occur at normalized frequencies (kb) given by the roots of

$$J_n'(kb) = 0, \tag{1}$$

where J_n are Bessel functions of order n. For each order "n" there exists a series of harmonics at frequencies $k_n^m b$, $m = 1 \rightarrow \infty$, each normal mode of vibration being denoted (n,m). At these frequencies the wave field in the cavity is of the form $J_n(k_n^m r)\cos(n\phi)$, where r, ϕ are the polar coordinates. For integrable geometries, such as a cavity of circular cross section, there is no level-repulsion and degeneracies can exist in the spectrum; i.e., different modes can occur at the same frequency. In addition, due to the circular symmetry, the majority of the modes of the circular cavity

FIG. 1. Scattering geometry.

FIG. 2. The schlieren visualization system.
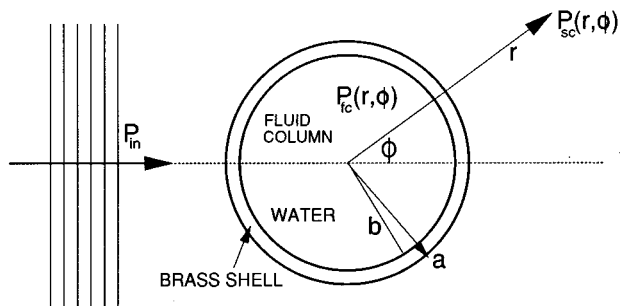
are doubly degenerate, splitting into two distinct modes when the shape of the cavity is perturbed.[11]

In real scattering systems, however, the interior wall of the shell presents a finite impedance to the interior acoustic field (as the boundary is not rigid) and the fluid column resonances have a finite width. In this case it is possible for neighboring resonances to overlap, provided their center frequencies are close together. This typically happens at high frequencies where the density of resonances is high and the resolution of individual modes is not possible. For metallic shells in water the resonances of the inner fluid column are given by the roots of

$$J_n'(kb) + i\,\frac{\rho_f c_f}{Z}\,J_n(kb) = 0, \tag{2}$$

where Z is the mechanical impedance of the shell and ρ_f and c_f are the density and velocity of the inner fluid, respectively.[11] For a high-impedance interface, such as the brass–water interface of interest in this paper, the roots of (2) are close to the roots of (1). From expression (2) Eq. (1) is obtained in the limit $Z \to \infty$.

In the current paper the resonances of a water-filled brass cylindrical shell are excited by insonifying the shell with a continuous plane wave. The scattering geometry is shown in Fig. 1. If the cylinder is insonified normal to its axis the problem is two dimensional and is easily modeled using the normal mode series approach.[14] The pressure in the fluid column is given by

$$P_{fc}(r,\phi) = P_0 \sum_{n=0}^{\infty} \epsilon_n i^n A_n(kb) J_n(kr)\cos(n\phi), \tag{3}$$

where $\epsilon_n = 1$ for $n=0$, $\epsilon_n = 2$ for $n>0$ and A_n are scattering coefficients determined by solving the boundary conditions at the inner and outer surfaces of the shell.[14] Equation (3) can be used to calculate the field in the fluid column at any frequency; both at, and away, from resonance. This formulation includes the elastic response of the shell and correctly accounts for the shell resonances and interfacial (Stoneley) waves, as well as the fluid column resonances. The frequencies at which the fluid column resonances occur can be estimated using Eq. (1) or (2), and then Eq. (3) can be used to calculate the interior field at frequencies in the vicinity of the estimated roots.
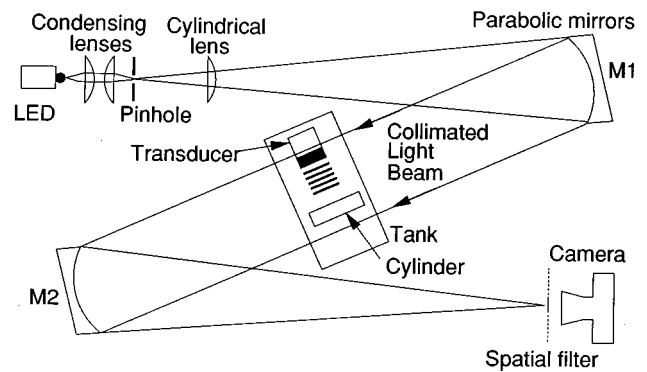
II. EXPERIMENTAL SYSTEM: SCHLIEREN

Light passing through an acoustic wave field in a fluid is refracted due to the density changes in the medium. If the intensity of the acoustic field is low and the diffraction effect is small the light suffers only phase variations and Raman–Nath diffraction results. In this case the acoustic field acts as a phase grating which produces an interference pattern containing information about the acoustic field. An ''image'' of the acoustic field is obtained by processing the information in the interference pattern. The relationship that this image has with the acoustic field depends upon the optical processing applied in the diffraction plane and the amplitude of the acoustic pressure field through which the light passes.[4,5,8]

The experimental system (Fig. 2) is based upon a ''Z'' configuration involving parabolic mirrors. 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 mirror (M1) and passed through the water-filled tank containing the acoustic field and cylindrical shell. The acoustic source 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 (M2) and forms a diffraction pattern in the focal plane of that mirror. 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. If a part of the diffraction pattern is removed by a spatial filter, and the remaining light allowed to recombine, a meaningful image is obtained on film or video. For very low acoustic pressures and central order filtering the schlieren image is proportional to the square of the acoustic pressure. At the higher acoustic pressures existing in the fluid column at resonance the pressure squared approximation breaks down, but a meaningful representation of the acoustic field is still obtained.[8]

The resonances of cylindrical shells are easily located, identified, and recorded using the schlieren system. The cylinder is placed in the water tank, correctly aligned and insonified by the transducer, and the frequency of the continuous-wave drive signal is swept. 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 sample, acoustic standing waves are clearly seen in the fluid

column, sometimes also in the outer fluid. At other frequencies, where little energy penetrates the shell, only the scattered field is observed. The imaging of resonances and interpretation of the images is described in greater detail in Ref. 8.

III. RESULTS AND DISCUSSION

Using the experimental system described above measurements have been made over a wide range of frequencies using a water-filled brass cylindrical shell, 120-mm long,

(a) 526.2 kHz ($kb = 32.14$)

(b) 567.4 kHz ($kb = 34.65$)

(c) 607.9 kHz ($kb = 37.13$)

(d) 663.8 kHz ($kb = 40.54$)

(e) 2.5 MHz ($kb = 152.7$)

(f) 3.5 MHz ($kb = 213.8$)
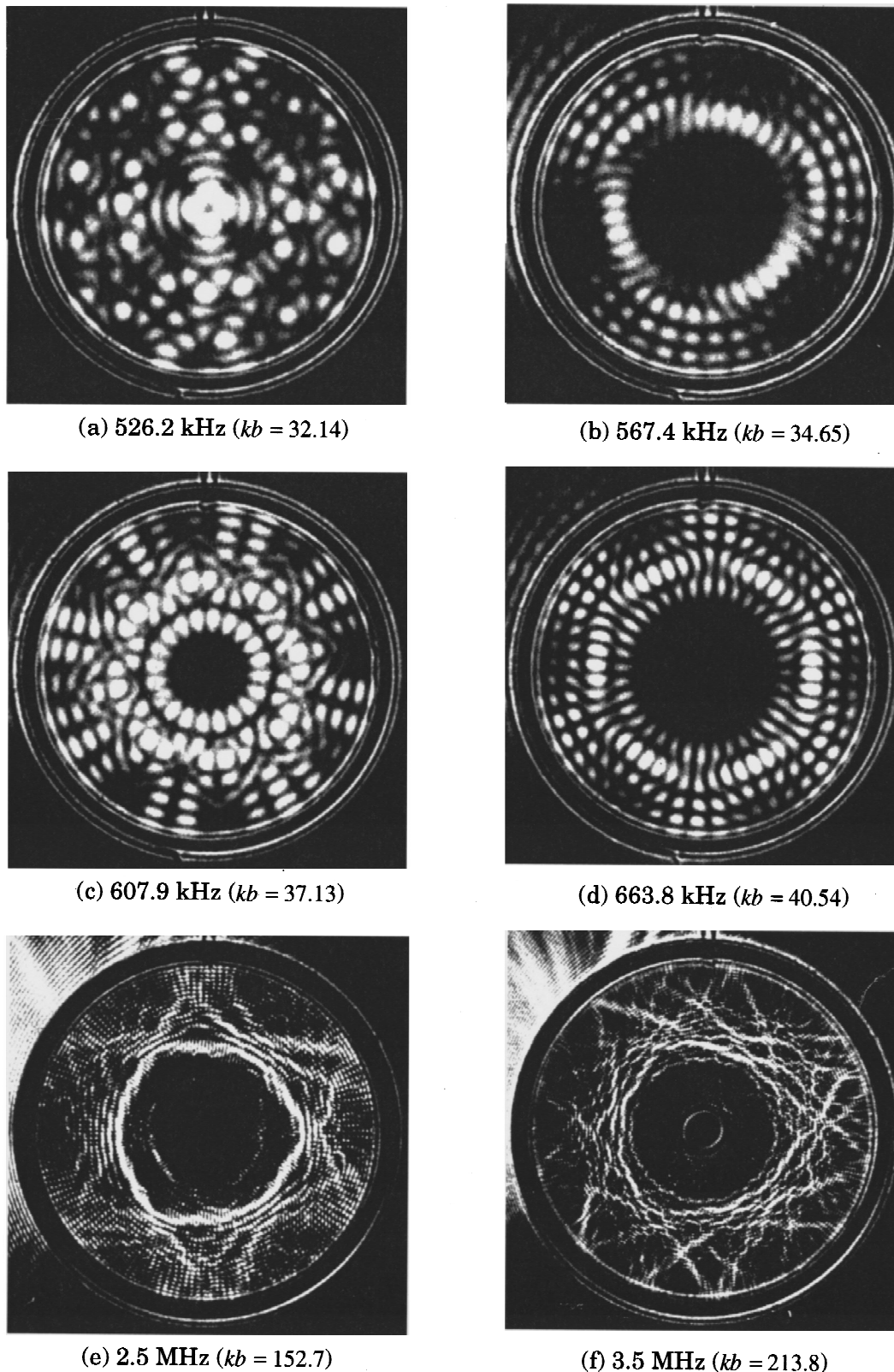
FIG. 3. High-frequency resonance modes of the fluid-filled cylindrical shell; experimental observations obtained with schlieren.

with outer and inner radii (a) 15.85 mm and (b) 14.25 mm ($b/a = 0.9$). Examples of the low-frequency resonance modes of such shells have been given in Refs. 8 and 10. In that low kb region individual modes were easily resolved and the familiar standing-wave patterns observed with schlieren. However, it is the higher frequency regime that is of interest in this paper.

At higher frequencies the density of resonances increases and many tend to overlap, making the resolution of individual resonances impossible. This effect can also be seen in the backscattered form function, and in calculations of the pressure in the fluid column itself. Indeed at very high frequencies (several megahertz for the shell considered here) the concept of a resonance as a single mode is not meaningful, as hybridization of modes becomes extensive. At the very highest frequencies *many* normal modes may contribute to the acoustic field at a given frequency.

The effect that this hybridization has upon the interior field is shown in the schlieren photographs presented in Fig. 3. New kinds of unexpected symmetry such as the threefold symmetry seen in Fig. 3(b) can appear in the resonance patterns. The dissymmetrization of the wave fields occurs due to the interference of overlapping resonances, resulting in a reduction in the symmetry of the whole in comparison with the symmetry of the individual modes. Resonances having other symmetries, such as fivefold and sevenfold symmetry have also been observed with the system. Theoretical calculations of the pressure fields within the fluid column are presented in Fig. 4 for several of the modes seen experimentally in Fig. 3. For these calculations the normal mode series solution for a plane wave normally incident upon an infinite cylindrical shell was used [Eq. (3)]. The agreement between theory and experiment is good; all the main features observed in the experiment are reproduced in the predictions. It should be noted that the experimental images differ in that the wave fields are rotated with respect to the theoretical calculations; this is most probably due to the presence of the supporting thread, or the insonifying beam.

The behavior manifest in Fig. 3(a)–(d) (and Fig. 4) can be explained as follows. At these frequencies, only two modes are likely to interfere, and the new symmetry seen in the acoustic field can be thought of as arising from a two-dimensional beating effect. If we just consider the angular parts of two overlapping wave fields, having n_1 and n_2 acoustic wavelengths around the circumference, the wave field resulting from a superposition would have an angular variation given by

$$\cos(n_1\phi) + \cos(n_2\phi) = 2\cos\left(\frac{n_1 + n_2}{2}\phi\right)\cos\left(\frac{n_1 - n_2}{2}\phi\right). \tag{4}$$

Take the field shown in Fig. 3(b) for example. At this frequency there exists two overlapping resonances having symmetries $n = 19$ and $n = 16$. Hence $19 - 16 = 3$ appears in the right-hand side of (4), giving a factor $\cos(3\phi/2)$ which results in the threefold behavior seen in Figs. 3(b) and 4(a). A plot of the above function (4) for $n_1 = 16$ and $n_2 = 19$ is given in Fig. 5. This angular variation agrees well with that of Fig. 4(a), including the differences in symmetry in the three main
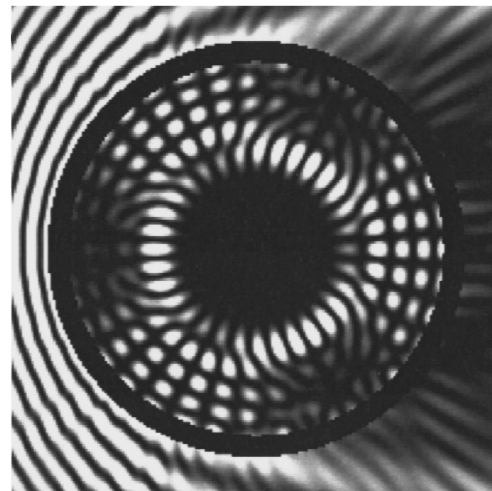
(a) $kb = 34.66$
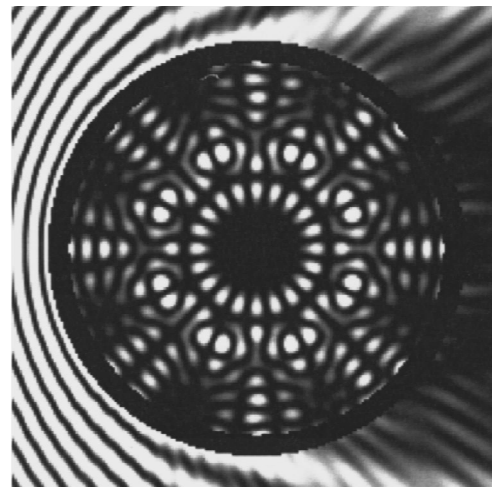
(b) $kb = 37.12$
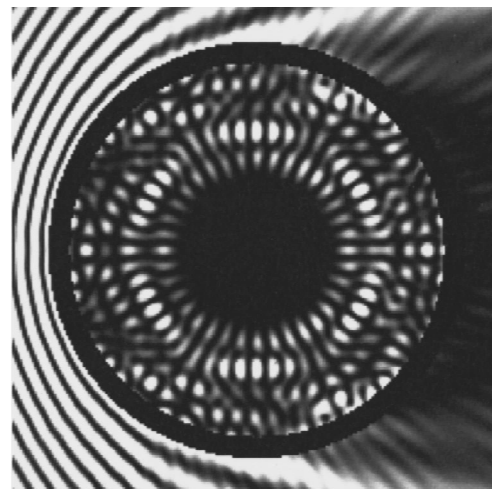
(c) $kb = 40.49$

FIG. 4. Theoretical predictions corresponding to Fig. 3(b), (c), and (d).

lobes. Similar analysis can be done for other hybrid resonances.

At higher frequencies many modes overlap and the concept of a resonance disappears. Examples of wave fields observed at these frequencies are shown in Fig. 3(e) and (f).
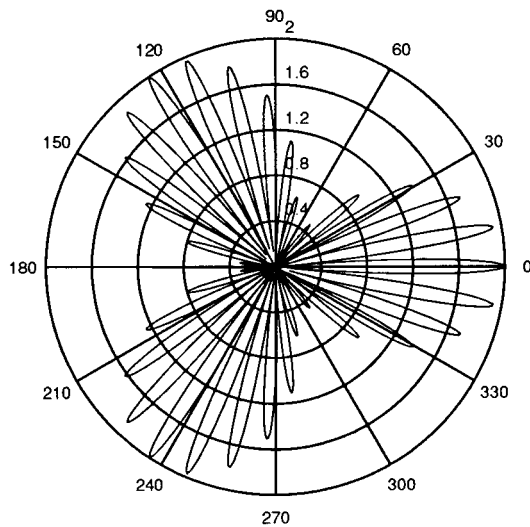
FIG. 5. Plot of $|\cos(n_1\theta)+\cos(n_2\theta)|$ for $n_1=16$, $n_2=19$ illustrating the beating effect for overlapping $n=16$ and $n=19$ modes, cf. Fig. 4(a).

FIG. 6. Ray paths (orbits) within a circular cavity; (a) an orbit that never repeats; a caustic is formed [cf. Fig. 3(f)]; (b) a nonisolated closed orbit of three bounces [cf. Fig. 3(b) and (e)] and an orbit of six bounces [cf. Fig. 3(d)].

The appearance of the images shown here are reminiscent of the classical ''ray'' or ''orbit'' concept that is applicable to wave fields in the high-frequency (short wavelength) domain. In particular, the association of some of the features of the resonance patterns with periodic orbits of a ray (or particle) ''bouncing'' around within the enclosure, can be made. In the circular cavity a typical orbit, such as that shown in Fig. 6(a), never repeats itself: every point on the circumference is ''hit'' eventually, and a caustic is formed. This kind of raylike behavior is most evident in the experimental images shown in Fig. 3(e) and (f).

The rather special cases where a ray repeatedly traverses the same (periodic) orbit are illustrated in Fig. 6(b) and (c). Some of the acoustic resonance fields presented in Fig. 3 can clearly be associated with this kind of behavior. The triangular orbit of Fig. 6(b) can be seen in Fig. 3(e), resulting in the Star of David pattern. Even the comparatively lower-frequency modes of Fig. 3(b) and (d) have a symmetry that can be associated with periodic orbits such as those illustrated in Fig. 6(b) and (c), respectively.

IV. CONCLUSION

The use of a schlieren visualization system for studying the high-frequency resonance behavior of two-dimensional fluid-filled cavities has been demonstrated, and results for the wave field inside the fluid column of a brass cylindrical shell have been presented. Hybridization of modes is seen to occur at frequencies where resonances overlap and the resulting wave fields exhibit dissymmetrization. For moderately high frequencies ($kb=30-40$) where only two modes overlap the observed phenomena are explained in terms of a two-dimensional beating effect. Theoretical calculations using the normal-mode series solution validate these experimental observations. At the very highest frequencies considered raylike behavior is observed and caustics are formed in the wave field in the fluid column.

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 geometry and is of particular value when considering those geometries that cannot easily be modeled. Of particular interest are the modes of cavities of noncircular geometry, such as elliptical-[11] and stadium-[13] shaped cavities.
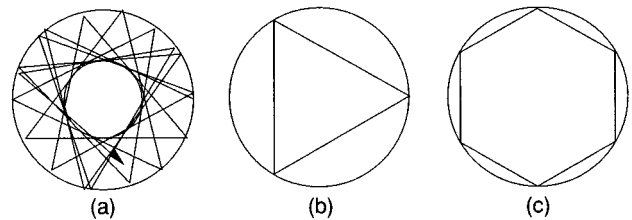
ACKNOWLEDGMENTS

[1] L. Flax, G. C. Gaunaurd, and H. Überall, ''Theory of resonance scattering,'' in *Physical Acoustics Vol. XV*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1981).

[2] G. Maze, J. L. Izbicki, and J. Ripoche, Ultrasonics **24**, 354–362 (1986).

[3] R. B. Barnes and C. J. Burton, ''Visual methods for studying ultrasonic phenomena,'' J. Appl. Phys. **20**, 286–294 (1949).

[4] J. A. Bucaro, L. Flax, H. D. Dardy, and W. E. Moore, ''Image profiles in schlieren observations of acoustic wavefronts,'' J. Acoust. Soc. Am. **60**, 1079–1084 (1976).

[5] J. A. Bucaro and H. D. Dardy, ''Sensitivity of the schlieren method for the visualization of low-frequency ultrasonic waves,'' J. Acoust. Soc. Am. **63**, 768–773 (1978).

[6] W. G. Neubauer, ''Observation of acoustic radiation from plane and curved surfaces,'' in *Physical Acoustics Vol. X*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1973), pp. 61–126.

[7] R. Reibold, ''Light diffraction tomography applied to the investigation of ultrasonic fields, Part II: standing waves,'' Acustica **63**, 283–289 (1987).

[8] P. A. Chinnery, V. F. Humphrey, and C. Beckett, ''The schlieren image of two-dimensional ultrasonic fields and cavity resonances,'' J. Acoust. Soc. Am. **101,** 250–256 (1997).

[9] S. M. Knapp and V. F. Humphrey, ''Schlieren visualisation of low frequency ultrasonic fields,'' in *Ultrasonics International 89 Conference Proceedings* (Butterworth, Washington, DC, 1989), pp. 1089–1094.

[10] 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.

[11] P. A. Chinnery and V. F. Humphrey, ''Fluid column resonances of water-filled cylindrical shells of elliptical cross section,'' J. Acoust. Soc. Am. (to be published); ''Resonances of deformed cylindrical shells—experimental visualisation and identification,'' in *Proceedings of the 2nd European Conference on Underwater Acoustics*, edited by L. Björnö (European Commission, 1994), pp. 159–164.

[12] P. A. Chinnery and V. F. Humphrey, ''Resonances in acoustic scattering by cylindrical objects,'' Proc. IOA **16**, 51–58 (1994).

[13] P. A. Chinnery and V. F. Humphrey, ''Experimental visualization of acoustic resonances within a stadium-shaped cavity,'' Phys. Rev. E **53**, 272–276 (1996).

[14] N. D. Veksler, *Resonance Acoustic Spectroscopy* (Springer-Verlag, Berlin, 1993).

Study on the longitudinal–torsional composite vibration of a sectional exponential horn

Lin Shuyu

Applied Acoustics Institute, Shaanxi Teachers University, Xian, Shaanxi, 710062, People's Republic of China

The longitudinal–torsional composite vibration of a sectional exponential horn was studied theoretically and experimentally. The horn consists of an exponential horn with an untapered rod at its large end. Based on the plane-wave approximation, the resonance frequency equations and other characteristics for the longitudinal and torsional vibrations of the horn were derived and some theoretical curves were given. Since the sound speeds of the longitudinal and torsional vibrations are different in the same material, the longitudinal and torsional vibrations in a horn are generally difficult to resonate at the same resonance frequency. In this paper, to overcome this problem, the cross-sectional radius decay coefficient of the exponential part in the sectional horn, which determines the sound speeds and other performance parameters of the horn, is selected to make the longitudinal and torsional vibrations resonate at the same frequency. By means of this method, the simultaneous resonance of the longitudinal and torsional vibrations in the horn is achieved by choosing the proper value of the radius decay coefficient. Some sectional exponential horns were designed and made, their resonance frequencies were measured, and the vibrational characteristics of the horn under large excitation were observed qualitatively. Experimental results show that the measured resonance frequencies of the horns are in good agreement with the theoretical results, and the measured resonance frequencies of the longitudinal and torsional vibrations are also in good agreement with each other. Therefore the simultaneous resonance of the longitudinal and torsional vibrations in the same horn are achieved both theoretically and experimentally. This kind of sectional exponential horn can be used in ultrasonic machining, ultrasonic drilling, ultrasonic fatigue testing, and other applications which need very large mechanical displacement amplitudes. © *1997 Acoustical Society of America.* [S0001-4966(97)05908-0]

PACS numbers: 43.20.Mv, 43.20.Ye, 43.35.Zc [JEG]

INTRODUCTION

Ultrasonic solid horns are important parts in high-power ultrasonic vibrating systems such as ultrasonic machining and soldering. In traditional ultrasonic applications, vibrational mode of the vibrating system is generally single, such as the longitudinal or torsional modes. The design theory for the ultrasonic horn with a single vibrational mode was well established.[1–4] With the development of some new ultrasonic technologies such as ultrasonic motor and ultrasonic revolving machining, composite vibrational modes such as longitudinal–torsional and longitudinal–flexural vibrations are needed.[5–7] As for the excitation of longitudinal–torsional composite vibration in a horn or a tool, the traditional method is to use two sets of longitudinal sandwiched ultrasonic transducers.[8] One set of transducers is attached on the large end of the horn, which produces the longitudinal vibration and the other set of transducers (usually there are two transducers) are attached on the opposite sides of the horn, which cause the horn to produce the torsional vibration. Using this kind of excitation method, the horn or tool will produce the longitudinal–torsional composite vibration. In general cases, however, the longitudinal and torsional vibrations in a horn resonate at different frequencies. Therefore to achieve the mechanical resonance of longitudinal and torsional vibrations in the horn, these two sets of longitudinal transducers must have different resonance frequencies. This

makes the fabrication of the transducers complex and the excitation of the transducers difficult. The cause is that the two sets of transducers have different frequency and impedance characteristics, and therefore, two sets of electric generators of different frequency and impedance characteristics are needed. However, if the ultrasonic horn or tool has the same resonance frequency of longitudinal and torsional vibrations, its excitation will become simple, and only one set of electric generators is needed. In this paper, the simultaneous resonance of longitudinal and torsional vibrations in the sectional horn was studied, and some experimental results were given. These can be used in the design and calculation of high-power ultrasonic vibrating systems.

I. THE SIMULTANEOUS RESONANCE OF LONGITUDINAL AND TORSIONAL VIBRATIONS IN THE SECTIONAL EXPONENTIAL HORN

In Fig. 1, the sectional exponential horn and the coordinate system are shown, where L_1 and L_2 are the lengths of the untapered and exponential parts, and S_1, R_1 and S_2, R_2 are the cross-sectional area and radius of the horn at its large and small ends. The variation law of cross-sectional area of the exponential part is $R(x)=R_1 \exp(-\beta X)$. Here, β is the radius decay coefficient of the horn. It is used to change the sound speeds of longitudinal and torsional vibrations in the horn and to make the longitudinal and torsional
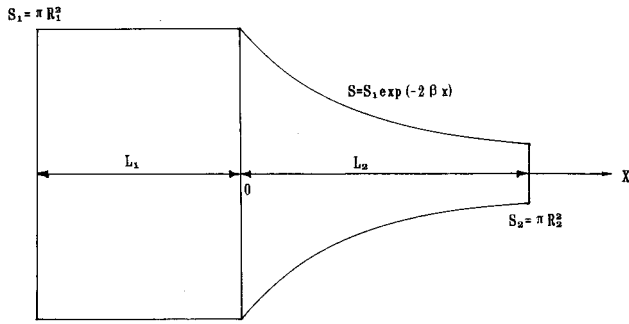
FIG. 1. Illustration of the sectional exponential horn and its coordinate system.

vibrations resonate at the same frequency. Based on the plane-wave approximation,[9,10] the vibration of the horn can be regarded as one dimensional. The resonance frequency equation of longitudinal vibration in the sectional horn is

$$\tan(k_L L_1) = - \frac{k_L L_2}{\beta L_2 + k_L' L_2 \, \cot(k_L' L_2)}, \qquad (1)$$

where $\beta L_2 = \ln N$, $N = R_1/R_2$, $k_L = \omega/c_L$ is the longitudinal wave number, $c_L = (E/\rho)^{1/2}$ is the sound speed of longitudinal vibration in an untapered slender rod, $k_L' = (k_L^2 - \beta^2)^{1/2} = \omega/c_L'$, and k_L' and c_L' are the wave number and sound speed of longitudinal vibration in exponential horn. The longitudinal displacement magnification M_L and input mechanical impedance Z_{iL} can be expressed as

$$M_L = \frac{N \cos(k_L L_1)}{\ln N \, \sin(k_L' L_2)/(k_L' L_2) + \cos(k_L' L_2)}, \qquad (2)$$

$$Z_{iL} = jZ_0^L \frac{k_L \tan(k_L' L_2) + \tan(k_L L_1)[k_L' + \beta \, \tan(k_L' L_2)]}{[k_L' + \beta \, \tan(k_L' L_2)] - k_L \, \tan(k_L' L_2)\tan(k_L L_1)}, \qquad (3)$$

where $Z_0^L = \rho c_L S_1$ is the characteristic impedance of the horn at its large end. It is obvious that the frequency equation (1) can be obtained from Eq. (3) by letting $Z_{iL} = 0$. The above equations can be used to design and calculate the sectional horn in longitudinal vibrational mode. For the torsional vibration, the resonance frequency equation of torsional vibration in the sectional horn is

$$\tan(k_t L_1) = - \frac{k_t L_2}{2\beta L_2 + k_t' L_2 \, \cot(k_t' L_2)}, \qquad (4)$$

where $k_t = \omega/c_t$ is the wave number of torsional vibration, $c_t = (G/\rho)^{1/2}$ is the sound speed of torsional vibration in a untapered slender rod, $k_t' = (k_t^2 - 4\beta^2)^{1/2} = \omega/c_t'$, and k_t' and c_t' are the wave number and sound speed of torsional vibration in the exponential horn. The angular displacement magnification and the input mechanical impedance of the horn in torsional vibration can be obtained as

$$M_T = \frac{N^2 \cos(k_t L_1)}{2 \ln N \, \sin(k_t' L_2)/(k_t' L_2) + \cos(k_t' L_2)}, \qquad (5)$$

$$Z_{it} = jZ_0^t \frac{k_t \tan(k_t' L_2) + \tan(k_t L_1)[k_t' + 2\beta \, \tan(k_t' L_2)]}{[k_t' + 2\beta \, \tan(k_t' L_2)] - k_t \, \tan(k_t' L_2)\tan(k_t L_1)}, \qquad (6)$$

where $Z_0^t = \rho c_t I_{p1}$ is the characteristic impedance of the horn in torsional vibration at its large end and $I_{p1} = \pi R_1^4/2$. Using the above equations, the sectional horn in torsional vibration can be designed.

From Eqs. (1) and (4) it can be seen that when the material and dimensions of the horn are given, the resonance frequencies of the horn in longitudinal and torsional vibrations can be obtained. However, in general cases, as the sound speeds of longitudinal and torsional vibrations in the horn are different, therefore, the longitudinal and torsional vibrations in the same horn are difficult to resonate at the same frequency. However, based on one-dimensional theory of longitudinal and torsional vibrations in a tapered slender rod, it can be seen that the sound speeds of longitudinal and torsional vibrations depend on the radius decay coefficient of the horn. Therefore, by choosing the radius decay coefficient of the horn, the sound speeds can be changed, and the longitudinal and torsional vibrations in the horn can be made to resonate at the same frequency. Using the theory of solid exponential horn, the sound speeds of longitudinal and torsional vibrations in the exponential horn can be expressed as

$$c_L' = \frac{c_L}{[1 - (\beta c_L/\omega)^2]^{1/2}}, \qquad (7)$$

$$c_t' = \frac{c_t}{[1 - (2\beta c_t/\omega)^2]^{1/2}}. \qquad (8)$$

It can be seen from Eqs. (7) and (8) that the sound speeds of longitudinal and torsional vibrations in an exponential horn depend on the radius decay coefficient of the horn. Furthermore, the changing speeds of the sound speeds of longitudinal and torsional vibrations with the radius decay coefficient are different. For a single exponential horn, the simultaneous resonance of longitudinal and torsional vibrations in the horn was achieved, and the analytic expressions for the radius decay coefficient and other parameters were given.[11] For the sectional exponential horn, since the changeable parameters are increased, their different combinations can satisfy the resonance frequency equations (1) and (4) easily.

Based on the above analysis, the longitudinal–torsional composite horn with the same resonance frequency of longitudinal and torsional vibrations can be calculated and fabricated using Eqs. (1) and (4). However, as the frequency equations are transcendental ones about the resonance frequency and the dimensions of the horn, the analytical solutions to these two equations are impossible to find. Therefore numerical methods must be used in the calculation of the longitudinal–torsional composite exponential horn. In this paper, one kind of numerical method is developed which can be used to compute and design the sectional exponential horn in longitudinal–torsional composite vibration. The procedure is as follows:

(a) Choose the material of the horn. For high-power ultrasonic applications, the material of the horn must be of low mechanical loss, large elasticity, high fatigue strength, and strong anticorrosion. According to these principles, titanium, duralumin, steel, and their alloys can be used.

(b) Determine the resonance frequency. For the longitudinal–torsional composite horn, there are two kinds
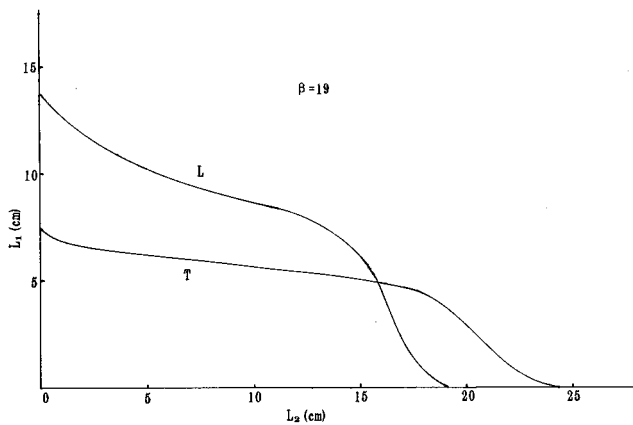
FIG. 2. Computed theoretical curve of the resonant length L_1 vs L_2 of the sectional exponential horn in longitudinal and torsional vibrations.

FIG. 3. Computed theoretical curve of the displacement magnification versus the resonant length L_2 of the horn in longitudinal and torsional vibrations.

of resonance frequency, one is for the longitudinal vibration, the other is for the torsional vibration. The objective of this study is to make these two frequencies equal to each other. Therefore when the horn is designed and fabricated, the resonance frequencies for longitudinal and torsional vibrations should be equal to each other, and they must be equal to the designed resonance frequency of the horn.

(c) When the material and the resonance frequency of the horn are given, there are three separate unknowns in Eqs. (1) and (4). Then, let the radius decay coefficient be a parameter, for a certain value of the length L_2 of the exponential part in the horn, two values of the length of the back untapered rod can be obtained from the frequency equations, i.e., L_{11} and L_{12}, these two lengths are the resonance lengths of the horn in longitudinal and torsional vibrations. In general cases, these two lengths are different.

(d) Change the value of the length L_2 of the exponential part in the horn until the two values of the length L_1 of the untapered rod in the horn from the frequency equations are equal to each other. In this case, the longitudinal and torsional vibrations in the horn will resonate at the same frequency which was designated in the procedure (b).

(e) Using the relation of $\beta L_2 = \ln N$, the area coefficient can be determined. According to different applications, the radius of the horn at its large end and small end can be obtained from the expression of $N = R_1/R_2$.

(f) Based on the equations derived in the above analysis, the displacement magnification of the horn in longitudinal and torsional vibrations can be computed.

From the above analysis, it can be seen that it is not until the procedure (e) that the theoretical design of the sectional exponential horn in longitudinal–torsional composite vibration is completed. Figures 2 and 3 are theoretical resonant and displacement magnification curves computed from the resonance frequency equations (1) and (4) and Eqs. (2) and (6). These two curves show the relation between the resonant lengths L_1 and L_2 of the horn and the relation between the magnification and the resonant length L_2 of the horn in longitudinal and torsional vibration. In Figs. 2 and 3 the radius decay coefficient of the horn is given a certain value. When the radius decay coefficient of the horn is changed, the resonant lengths and the magnification will change accordingly.
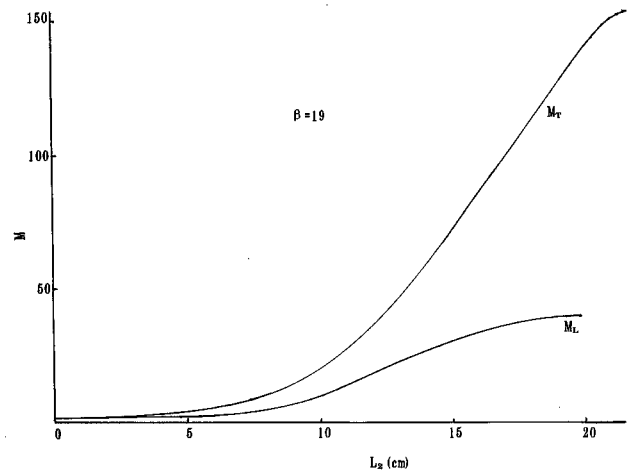
In the figures, L and T represent the longitudinal and torsional vibration, M_L and M_T are the displacement magnification of the horn in longitudinal and torsional vibrations.

It can be seen from Figs. 2 and 3 that when the dimensions of the horn satisfy the resonance frequency equations (1) and (4), the longitudinal and torsional vibrations resonate simultaneously at the same frequency. This corresponds to the intersecting point in Fig. 2. Since the resonance frequency equations are transcendental ones, the analytic expressions for the resonant length and the radius decay coefficient at this intersecting point are impossible to find. However, since there are four changeable parameters $(L_1, L_2, \beta, \omega)$ in the resonance frequency equations, this intersecting point in Fig. 2 can be changed, i.e., the simultaneous resonance of the sectional horn with different dimensions can be achieved by changing the radius decay coefficient and the length of the horn. However, this is different from the single exponential horn. For a single exponential horn, when the longitudinal and torsional vibrations resonate at the same frequency, the sound speeds of longitudinal and torsional vibrations are equal to each other. While for the sectional exponential horn, since the length L_1 also affects the resonance frequency, the sound speeds for the longitudinal and torsional vibrations are different when the simultaneous resonance was achieved.

From the above analysis and the curves in Figs. 2 and 3, it can be seen that the radius decay coefficient of the horn affects the performances of the sectional horn. It is obvious that the larger the radius decay coefficient, the larger the magnification of the horn. Therefore the radius decay coefficient of the exponential horn can be used to optimize some performance parameters of the horn, such as the displacement magnification and the input mechanical impedance. As for which parameter is optimized, it depends on the actual applications that need different performance requirements. On the other hand, it can be seen from Figs. 2 and 3 that the magnification of longitudinal displacement is smaller than that of torsional angular displacement of the horn, and the resonance frequency change arising from the change of the resonant length of the horn in torsional vibration is larger
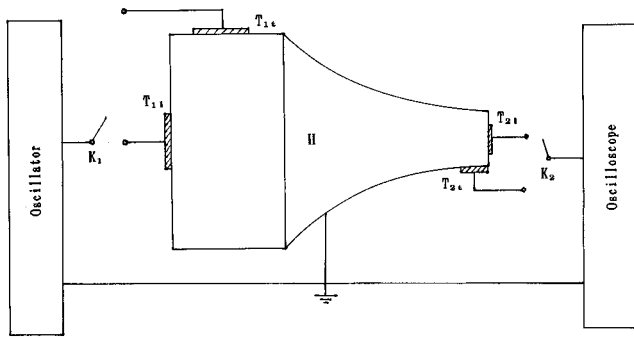
FIG. 4. Setup for measuring the resonance frequencies of the horns.

than that in longitudinal vibration. These can give us some hints in the design and fabrication of the longitudinal–torsional composite vibration horn, especially in the cases that the length of the horn changes such as in ultrasonic machining and drilling.

II. EXPERIMENTS

In the above analysis, the longitudinal–torsional composite sectional exponential horn was studied theoretically and the simultaneous resonance of longitudinal and torsional vibrations in the horn was achieved. To study the characteristics and the simultaneous resonance of the horn experimentally, some sectional exponential horns were fabricated according to the frequency equations derived in the above analysis. The material of the horn is steel, its material parameters are as follows: $\rho = 7.8 \times 10^3$ kg/m^3, $E = 1.95 \times 10^{11}$ N/m^2, $G = 7.62 \times 10^{10}$ N/m^2, $\sigma = 0.28$. The experiment was divided into two parts, one is the measurement of the resonance frequencies under small excitation, the other is the high-power excitation of the horn and the measurement of the resonance frequencies. In the second part of the experiment, the vibration characteristics of the horn under high-power excitation was also studied qualitatively. The experiment was described in the following.

A. The measurement of resonance frequencies of the sectional exponential horns

Under small excitation, the resonance frequencies of the horn were measured using the emitting–receiving method. The experimental setup is illustrated in Fig. 4, where K_1 and K_2 are two switches. When K_1 and K_2 are switched on T_{1L} and T_{2L}, the resonance frequencies of the horn in longitudinal vibration were measured. When K_1 and K_2 are switched on T_{1t} and T_{2t}, the resonance frequencies of the horn in torsional vibration were measured. In Fig. 4, T_{1L} and T_{2L} are two piezoelectric ceramic circular transducers, their vibrational modes are thickness extensional vibration. T_{1t} and T_{2t} are two piezoelectric ceramic rectangular transducers

that can produce the thickness shearing vibration. The geometrical dimensions of the measuring transducers are small so that their resonance frequencies are much higher than the resonance frequencies of the horns to be measured, therefore, the measuring transducers have little effect on the measurement of the resonance frequencies of the horn. There are two steps in the measurement of the resonance frequencies of the horns. First, K_1 and K_2 are switched on T_{1L} and T_{2L}, change the frequency of the oscillator until the millivolt meter has a maximum, the frequency corresponding to the maximum reading is the longitudinal resonance frequency of the horn. Second, K_1 and K_2 are switched on T_{1t} and T_{2t}, repeat the same procedure as that described in the first step, the torsional resonance frequency of the horn can be measured. The measured results are shown in Table I, where f is the designed frequency of the horn, f_{Lm} and f_{tm} are the measured resonance frequencies of the horn in longitudinal and torsional vibrations, respectively. It can be seen from Table I that the measured resonance frequencies of the horn are in good agreement with the designed frequency, and the measured resonance frequencies of the horn in longitudinal and torsional vibrations are also in good agreement with each other. Therefore the simultaneous resonance of longitudinal and torsional vibrations in the horn was achieved experimentally.

As for the error for the measured resonance frequencies, the following factors should be taken into account. First, the standard values of the material parameters used in the computation of the resonance frequency are different from the actual values of the material parameters. Second, in order to satisfy the resonance frequency equations and to achieve the simultaneous resonance, the radius decay coefficient is generally large, this means that the cross-sectional area of the exponential horn changes greatly. This may cause some frequency error and the measured frequencies are smaller than the predicted ones. However, this error is small and negligible, since the lateral dimensions are smaller than a quarter of the wavelength in the horn. Third, in the experiment, the excitation of the torsional vibration is by means of the thickness shearing vibration of rectangular piezoelectric ceramics. In this case, the torsional vibration is not an ideal circumferential vibration.

B. The high-power excitation of the longitudinal–torsional horn and the resonance frequency measurement of the horn under high-power excitation

In practical applications, such as ultrasonic machining and drilling, the horn is usually excited by a large signal. For the high-power excitation of the longitudinal–torsional horn, there are two methods. One is to use two sets of longitudinal sandwiched transducers. In this method, the longitudinal

TABLE I. The measured resonance frequencies of the horns in longitudinal–torsional composite vibration.

R_1 (mm)	R_2 (mm)	L_1 (mm)	L_2 (mm)	β (m^{-1})	f (Hz)	f_{Lm} (Hz)	f_{tm} (Hz)
33.8	2.0	70.5	201.8	14.0	15 000	14 526	14 489
32.5	2.0	50.8	150.7	18.5	20 000	19 408	19 445

transducer attached on the large end of the horn causes the horn to produce longitudinal vibration, while the longitudinal transducers attached on the opposite sides of the horn cause the the the horn to produce the torsional vibration. Therefore the torsional vibration in the horn excited by this method is produced by conversion of longitudinal into torsional vibration. However, since the horn has the same resonance frequency of longitudinal and torsional vibrations, therefore, the exciting sandwiched transducers are all the same in both the frequency and impedance characteristics. The second method to excite the longitudinal–torsional horn is to use the longitudinal–torsional composite sandwiched transducer. This kind of transducers can produce both longitudinal and torsional vibrations. In the transducer, there are two sets of piezoelectric ceramic rings, one set of elements are longitudinally polarized piezoelectric ceramic stacks, which produce the longitudinal vibration, the other set of elements are tangentially polarized piezoelectric ceramic stacks, which produce the torsional vibration. The resonance frequencies of the transducer in longitudinal and torsional vibrations are equal to each other. The longitudinal–torsional transducer is attached on the large end of the horn. When the transducer is excited by an electric signal, it will produce the longitudinal and torsional vibrations, and at the same time, the longitudinal and torsional vibrations are also produced in the horn. Compared with the first method to excite the longitudinal–torsional horn, in the second method, only one transducer is needed, and the construction of the vibrating system is simple. However, the design and fabrication of the longitudinal–torsional sandwiched transducer is more complex than that of the sandwiched longitudinal transducer. As for the design and fabrication of the longitudinal–torsional transducer and the simultaneous resonance of longitudinal and torsional vibrations in the transducer, since this work is on the study of longitudinal–torsional sectional horn, they will be studied in later work. Although the vibrating system to excite the longitudinal–torsional horn used in the first method is complex in construction and big in volume, the sandwiched longitudinal transducers used are easy to design and fabricate, and the electric excitation of the transducers are also simpler than that of the longitudinal–torsional composite transducer. Therefore in the following study, the longitudinal–torsional sectional horn is excited by the first method, and the resonance frequencies of the vibrating system (including the longitudinal transducer and the horn) under large exciting signal were measured. The experimental setup is shown in Fig. 5, where BLT represents the longitudinal sandwiched transducer. The experimental results are shown in Table II, where P is the input electric power of the longitudinal transducers, f_T and f_H are the designed resonance frequencies of the sandwiched transducer and the sectional exponential horn, f_{TM} is the measured resonance frequencies of the transducer under small excitation using the transmission line method, f_{HL} and f_{HT} are the measured resonance frequencies of the horn in longitudinal and torsional vibrations under small excitation, f_1 is the measured resonance frequency of the vibrating system which consists of the horn and the longitudinal transducer attached on the large end of the horn under large excitation, f_2 is the measured
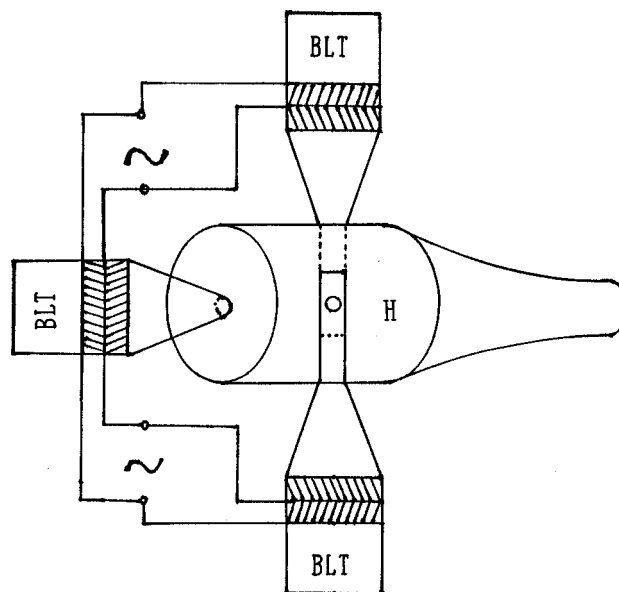
FIG. 5. Experimental setup for measuring the vibration characteristics of the sectional exponential horns under large excitation.

resonance frequency of the vibrating system which consists of the horn and the transducers attached on the opposite sides of the horn under large excitation, f_3 is the measured resonance frequency of the vibrating system which includes the horn, the longitudinal transducer attached on the large end of the horn, and the longitudinal transducers attached on the opposite sides of the horn also under large excitation. It can be seen from Table II that under large excitation, the measured resonance frequencies of the vibrating system in longitudinal and torsional vibrations are also in good agreement with each other, and the composite resonance frequencies of the system are basically consistent with the resonance frequencies of the system in single vibrational mode. On the other hand, it can also be seen that when the input electric power of the longitudinal transducer is increased, the resonance frequency of the vibrating system is decreased. The cause of this kind of frequency decrease is probably that when the input electric power is increased, the vibration of the system becomes intense, and the linking between the horn and the exciting transducers becomes loose, since all elements in the system are linked by screw botts. On the other hand, the heating of the system under large excitation also lowers the resonance frequencies. In this experiment, apart from measuring the resonance frequencies of the system, the qualitative observation of the vibration of the system was also carried out. When the system resonates, we can feel that the horn vibrates intensely. It is by means of this method that the resonance frequencies of the system under large excitation are measured. On the other hand, when the small end of the horn is immersed in water, streamings whirling about the tip of the horn can be observed, this is caused by the torsional vibration of the horn. At the same time, intense streaming out of the tip of the horn proves that the horn produces intense longitudinal vibration, therefore it is illustrated experimentally that the horn can produce longitudinal and torsional vibrations at the same time, and the

TABLE II. The measured results of the longitudinal–torsional horns under large excitation.

P (W)	f_T (Hz)	f_H (Hz)	f_{TM} (Hz)	f_{HL} (Hz)	f_{HT} (Hz)	f_1 (Hz)	f_2 (Hz)	f_3 (Hz)
50	20 000	20 000	19 877	19 408	19 445	19 213	19 301	19 111
100	20 000	20 000	19 877	19 408	19 445	19 175	19 048	18 892

simultaneous resonance of longitudinal and torsional vibrations is achieved in the sectional exponential horn.

III. CONCLUSIONS

In this paper, the sectional exponential horn in longitudinal and torsional vibrations was studied, the simultaneous resonance of longitudinal and torsional vibrations in the horn was achieved theoretically and experimentally. To sum up the above analysis, the following conclusions can be obtained.

(a) In genaral cases, since the sound speeds of longitudinal and torsional vibrations in the same material are different, the longitudinal and torsional vibrations in a horn are difficult to resonate at the same frequency. However, for exponential horns, by changing the radius decay coefficient of the exponential horn, the sound speeds can be changed, and the simultaneous resonance at the same frequency of longitudinal and torsional vibrations can be achieved.

(b) In a longitudinal–torsional composite sectional exponential horn, the displacement magnification of longitudinal vibration is smaller than that of torsional vibration, and the resonance frequency of the horn in torsional vibration is more easily affected by the length of the horn than that in longitudinal vibration.

(c) In practical applications, the longitudinal–torsional horns can be made to vibrate in two methods. The first method is to use longitudinal sandwiched transducers, in this method, the torsional vibration is produced by conversion of longitudinal vibration. Although the construction is complex, the design and fabrication of the longitudinal transducer are simple. The second method to excite the horn is to use a longitudinal–torsional composite transducer, this kind of transducer can produce both longitudinal and torsional vibration, and its longitudinal and torsional vibrations resonate at the same frequency. In the second method, although the construction of the vibrating system is simple, the design and fabrication of the longitudinal–torsional composite transducer are complex.

(d) This kind of longitudinal–torsional sectional exponential horn can be used in ultrasonic machining, ultrasonic drilling, ultrasonic fatigue testing, and other applications which need very large displacement amplitudes.

ACKNOWLEDGMENTS

[1] E. A. Neppiras, ''Mechanical transformer for producing very large motion,'' Acustica **13**, 368–370 (1963).
[2] E. Eisner, ''Design of sonic amplitude transformers of high magnification,'' J. Acoust. Soc. Am. **35**, 1367 (1963).
[3] E. Eisner and J. S. Seager, ''A longitudinally resonant stub for vibrations of large amplitude,'' Ultrasonics **3**, 88–98 (1965).
[4] E. A. Neppiras, ''Very high energy ultrasonics,'' Br. J. Appl. Phys. **11**, 143 (1960).
[5] S. Ueha, H. Nagashima, and M. Masuda, ''Longitudinal-torsional composite transducer and its applications,'' Jpn. J. Appl. Phys., Suppl. **26-2**, 188–190 (1987).
[6] H. Nagashima and S. Ueha, ''Study of longitudinal-torsional composite vibrator system,'' J. Acoust. Soc. Jpn. **45**, 94–100 (1989).
[7] K. Nakamura, M. Kurosawa, and S. Ueha, ''Characteristics of a hybrid transducer-type ultrasonic motor,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 188–193 (1991).
[8] L. D. Rozenberg, *Sources of High-Intensity Ultrasound* (Plenum, New York, 1968), Vol. 22.
[9] R. W. Pyle, Jr., ''Solid torsional horns,'' J. Acoust. Soc. Am. **41**, 1147–1156 (1967).
[10] A. V. Kharitonov, ''Torsional ultrasonic concentrators,'' Sov. Phys. Acoust. **7**, 310,311 (1962).
[11] L. Shuyu, ''Study on the longitudinal-torsional composite mode exponential ultrasonic horns,'' Ultrasonics **34**, 757–762 (1996).

An improved acoustic Fourier boundary element method formulation using fast Fourier transform integration

A. H. W. M. Kuijpers,[a] G. Verbeek, and J. W. Verheij

Computational and Experimental Mechanics, Faculty of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Effective use of the Fourier series boundary element method (FBEM) for everyday applications is hindered by the significant numerical problems that have to be overcome for its implementation. In the FBEM formulation for acoustics, some integrals over the angle of revolution arise, which need to be calculated for every Fourier term. These integrals were formerly treated for each Fourier term separately. In this paper a new method is proposed to calculate these integrals using fast Fourier transform techniques. The advantage of this integration method is that the integrals are simultaneously computed for all Fourier terms in the boundary element formulation. The improved efficiency of the method compared to a Gaussian quadrature based integration algorithm is illustrated by some example calculations. The proposed method is not only usable for acoustic problems in particular, but for Fourier BEM in general. © *1997 Acoustical Society of America.* [S0001-4966(97)04908-4]

INTRODUCTION

The use of the boundary element method (BEM) in acoustics, tailored to problems involving axisymmetric bodies with arbitrary boundary conditions, has received some attention in the literature the past decade. Conceptually, the acoustic BEM for axisymmetric bodies applies a Fourier series expansion of the angular dependency of the acoustic variables in the problem. As a result, the surface integral in the Kirchhoff–Helmholtz integral equation reduces to a line integral and an integral over the angle of revolution (circumferential integral). The advantages of this so-called Fourier BEM approach are evident. Discretization of the body requires only meshing of the generator of the body with line elements. Also, the cost of numerically solving the system of equations is reduced because of a substantial decrease in the number of unknowns. The drawback of this method is the increased complexity of both the mathematical formulation and numerical implementation of the method.

The computation of the circumferential integrals (around the symmetry axis of the body) causes considerable numerical problems for the implementation of the Fourier boundary element method. The integrand of these integrals can be singular and has an oscillatory nature. Numerical values for these integrals were obtained using Trapezium quadrature by Akyol.[1] This method provided accurate results, but it was pointed out that the efficiency of the integral computation needed further investigation. A different method for the computation of the integrals was proposed by Soenarko[2] and Juhl,[3] who reformulated the integrand and employed an (unspecified) series of elliptic integrals for the singular part of the integral and Gaussian quadrature for the regular part.

The evaluation of the circumferential integrals is time consuming and has to be done for each Fourier mode number

that is present in the boundary conditions for the acoustic problem. This hampers the effective application of the Fourier boundary element method for practical acoustic applications. A faster evaluation method proposed in this article is based on fast Fourier transform (FFT). This method handles the calculation of the integrals both efficiently and accurately for all Fourier harmonics simultaneously. It was implemented and tested and it compared favorably to an integration method based on Gaussian quadrature.

I. FOURIER BOUNDARY INTEGRAL EQUATIONS FOR ACOUSTICS

The Kirchhoff–Helmholtz integral equation is a mathematical description for the acoustic radiation of structures. Consider a simple axisymmetric body B (Fig. 1). The Kirchhoff–Helmholtz integral equation for the pressure $p(\mathbf{x})$ at an observer point $\mathbf{x}$ can be written as[4]

$$C(\mathbf{x}) \cdot p(\mathbf{x}) = \int_S \left[p(\mathbf{y}) \frac{\partial G(\mathbf{x},\mathbf{y})}{\partial \nu} - G(\mathbf{x},\mathbf{y}) \frac{\partial p(\mathbf{y})}{\partial \nu} \right] dS(\mathbf{y}),$$
(1)

with

$$G(\mathbf{x},\mathbf{y}) = \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})}$$
(2)

as the three-dimensional free-space Green's function, and $C(\mathbf{x})$ as a coefficient depending on the position of $\mathbf{x}$:

$$C(\mathbf{x}) = \begin{cases} 0, & \text{for } \mathbf{x} \text{ outside the acoustic medium,} \\ 1, & \text{for } \mathbf{x} \text{ inside the acoustic medium,} \\ \frac{1}{2}, & \text{for } \mathbf{x} \text{ on the smooth surface } S \text{ of the} \\ & \text{acoustic medium.} \end{cases}$$
(3)
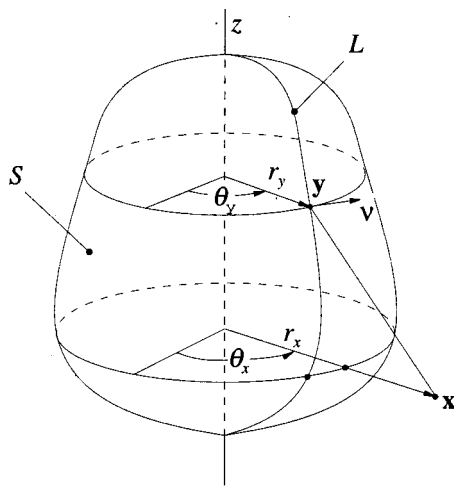
S: surface
L: body generator
$\mathbf{x}$: observer point with cylindrical
　　coordinates (r_x, θ_x, z_x)
$\mathbf{y}$: surface point with cylindrical
　　coordinates (r_y, θ_y, z_y)
ν: surface normal

FIG. 1. Simple axisymmetric body B.

For any other position of $\mathbf{x}$ on S for which there is no unique surface normal, for example when $\mathbf{x}$ is on an edge or a corner, the value of $C(\mathbf{x})$ is given by

$$
C(\mathbf{x}) = \begin{cases} 1 + \dfrac{1}{4\pi} \displaystyle\int_S \dfrac{\partial}{\partial \nu}\left(\dfrac{1}{R(\mathbf{x},\mathbf{y})}\right) dS(\mathbf{y}) \\ \quad \text{for the exterior acoustic problem,} \\[2mm] -\dfrac{1}{4\pi} \displaystyle\int_S \dfrac{\partial}{\partial \nu}\left(\dfrac{1}{R(\mathbf{x},\mathbf{y})}\right) dS(\mathbf{y}) \\ \quad \text{for the interior acoustic problem.} \end{cases}
$$
$$(4)$$

The distance R between the points $\mathbf{x}$ and $\mathbf{y}$ is defined as $R(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|$. The surface normal direction ν is directed inward the acoustic medium (see Fig. 1).

　　The geometry of an axisymmetric body with arbitrary boundary conditions can be described using a cylindrical coordinate system (r,θ,z). All variables then become functions of the cylindrical coordinates r, θ, and z, i.e.,

$$p(\mathbf{x}) = p(r_x, \theta_x, z_x),$$

$$p(\mathbf{y}) = p(r_y, \theta_y, z_y),$$

$$G(\mathbf{x},\mathbf{y}) = \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} = G(r_x, \theta_x, z_x; r_y, \theta_y, z_y),$$

$$dS(\mathbf{y}) = r_y \, d\theta_y \, dL_y,$$

with (r_x, θ_x, z_x) and (r_y, θ_y, z_y) as the coordinates of the observer point $\mathbf{x}$ and the surface point $\mathbf{y}$, respectively, and dL_y as the differential length of the generator L of the body at $\mathbf{y}$. Because of the axisymmetric properties of the body B the variables can be expanded in Fourier series. This expansion was reported by Soenarko[2] as follows:

$$p(\mathbf{y}) = \sum_{n=0}^{\infty} [p_n^s \sin(n\theta_y) + p_n^c \cos(n\theta_y)], \quad (5)$$

$$p(\mathbf{x}) = \sum_{n=0}^{\infty} [p_n^{s*} \sin(n\theta_x) + p_n^{c*} \cos(n\theta_x)], \quad (6)$$

with superscript $*$ to discern the Fourier coefficients for the surface point $\mathbf{y}$ and observer point $\mathbf{x}$. Note that the Fourier coefficients still depend on the coordinates r and z, but the dependence of pressure p on coordinate θ is expressed through the sine and cosine terms of the Fourier expansion. The other functions of Eq. (1) can be expanded likewise:

$$G(\mathbf{x},\mathbf{y}) = \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} = \sum_{n=0}^{\infty} [K_n^s \sin(n\theta_y)$$
$$+ K_n^c \cos(n\theta_y)], \quad (7)$$

$$\frac{\partial p}{\partial \nu}(\mathbf{y}) \equiv p'(\mathbf{y}) = \sum_{n=0}^{\infty} [p_n'^s \sin(n\theta_y) + p_n'^c \cos(n\theta_y)], \quad (8)$$

$$\frac{\partial G}{\partial \nu}(\mathbf{x},\mathbf{y}) \equiv G'(\mathbf{x},\mathbf{y}) = \sum_{n=0}^{\infty} [K_n'^s \sin(n\theta_y)$$
$$+ K_n'^c \cos(n\theta_y)]. \quad (9)$$

The Fourier coefficients in these equations are independent of θ_y but still dependent on the r_y and z_y. Observe that the Fourier coefficients of the expansions of the Green's function and its normal derivative (i.e., K_n^s, K_n^c, $K_n'^s$, and $K_n'^c$) are also dependent on all the cylindrical coordinates of point $\mathbf{x}$: r_x, θ_x, and z_x.

　　With the Fourier series description for the circumferential dependence of the acoustic variables, a modified form of the Kirchhoff–Helmholtz integral equation can be obtained. To that end, it is convenient to reformulate the Fourier coefficients of the Green's function and its derivative. The coefficients are determined by the standard Fourier transformation rules, for instance,

$$K_n^s = \frac{1}{\pi} \int_0^{2\pi} \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} \sin(n\theta_y) d\theta_y, \quad n = 0,1,2,\dots. \quad (10)$$

By defining $\theta \equiv \theta_y - \theta_x$ such that $d\theta = d\theta_y$ Eq. (10) can be written as

$$K_n^s = \frac{1}{\pi} \int_0^{2\pi} \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} \sin(n(\theta + \theta_x)) d\theta. \quad (11)$$

With the aid of a trigonometric identity, Eq. (11) can be rewritten as

$$K_n^s = \frac{1}{\pi} \int_0^{2\pi} \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} \sin(n\theta)\cos(n\theta_x) d\theta$$
$$+ \frac{1}{\pi} \int_0^{2\pi} \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} \cos(n\theta)\sin(n\theta_x) d\theta. \quad (12)$$

Because $\sin(n\theta)$ is an odd function of $\theta = \pi$ and the remainder of the integrand is symmetric around $\theta = \pi$ in the interval

[0,2π], the first integral of Eq. (12) vanishes. Introducing

$$H_n = \int_0^{2\pi} \frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} \cos(n\theta) d\theta, \tag{13}$$

Eq. (12) becomes

$$K_n^s = \frac{1}{\pi} H_n \sin(n\theta_x). \tag{14}$$

The cosine coefficients K_n^c of Eq. (7) can be derived in a similar manner:

$$K_n^c = \frac{1}{\pi} H_n \cos(n\theta_x). \tag{15}$$

The Fourier coefficients of Eq. (9) can also be determined analogously. Using

$$H_n' = \int_0^{2\pi} \frac{\partial}{\partial\nu} \left(\frac{e^{-ikR(\mathbf{x},\mathbf{y})}}{4\pi R(\mathbf{x},\mathbf{y})} \right) \cos(n\theta) d\theta, \tag{16}$$

they can be written as

$$K_n'^s = \frac{1}{\pi} H_n' \sin(\theta_x), \tag{17}$$

$$K_n'^c = \frac{1}{\pi} H_n' \cos(\theta_x). \tag{18}$$

The Fourier coefficients in Eqs. (5), (6), and (8) can be expressed similarly, for instance,

$$p_n^s(\mathbf{y}) = \frac{1}{\pi} \int_0^{2\pi} p(\mathbf{y}) \sin(n\theta_y) d\theta_y, \tag{19}$$

and similar expressions for $p_n^c(\mathbf{y})$, $p_n^{s*}(\mathbf{x})$, $p_n^{c*}(\mathbf{x})$, $p_n'^s(\mathbf{y})$, and $p_n'^c(\mathbf{y})$. By expanding all variables in Fourier series, Eq. (1) takes the form (where some of the summations take the index m for clarity)

$$C(\mathbf{x}) \left\{ \sum_{m=0}^{\infty} [p_m^{s*} \sin(m\theta_x) + p_m^{c*} \cos(m\theta_x)] \right\}$$

$$= \int_L \int_0^{2\pi} \left\{ \sum_{n=0}^{\infty} [p_n^s \sin(n\theta_y) + p_n^c \cos(n\theta_y)] \right\} \times \left\{ \sum_{m=0}^{\infty} \frac{1}{\pi} H_m'[\sin(m\theta_x)\sin(m\theta_y) + \cos(m\theta_x)\cos(m\theta_y)] \right\} r_y d\theta_y dL_y$$

$$- \int_L \int_0^{2\pi} \left\{ \sum_{n=0}^{\infty} [p_n'^s \sin(n\theta_y) + p_n'^c \cos(n\theta_y)] \right\} \times \left\{ \sum_{m=0}^{\infty} \frac{1}{\pi} H_m[\sin(m\theta_x)\sin(m\theta_y) + \cos(m\theta_x)\cos(m\theta_y)] \right\} r_y d\theta_y dL_y. \tag{20}$$

Matching the terms on the left- and right-hand sides of Eq. (20) and using the orthogonality properties for integrals involving $\sin(m\theta_y)\sin(n\theta_y)$, $\sin(m\theta_y)\cos(n\theta_y)$, and $\cos(m\theta_y) \times \cos(n\theta_y)$, the following expressions can be obtained after integration over $d\theta_y$:

$$C(\mathbf{x})p_n^{s*}(\mathbf{x}) = \int_L [p_n^s H_n' - p_n'^s H_n] r_y dL_y, \tag{21}$$

$$C(\mathbf{x})p_n^{c*}(\mathbf{x}) = \int_L [p_n^c H_n' - p_n'^c H_n] r_y dL_y. \tag{22}$$

All functions in Eqs. (21) and (22) are no longer explicitly dependent on angle θ. When the Fourier coefficients of surface pressure p_n^s, p_n^c and its normal derivative $p_n'^s$, $p_n'^c$ are known, the acoustic pressure at any observer point $\mathbf{x}$ inside, outside or on body B can be expressed as [see Eq. (6)]

$$C(\mathbf{x})p^*(\mathbf{x}) = \int_L \left\{ \sum_{n=0}^{\infty} [p_n^s H_n' - p_n'^s H_n] \sin(n\theta_x) \right.$$

$$\left. + \sum_{n=0}^{\infty} [p_n^c H_n' - p_n'^c H_n] \cos(n\theta_x) \right\} r_y dL_y. \tag{23}$$

When observer point $\mathbf{x}$ is on the surface of body B, Eqs. (21) and (22) can be rewritten as

$$C(\mathbf{x})p_n^s(\mathbf{x}) = \int_L [p_n^s H_n' - p_n'^s H_n] r_y dL_y, \tag{24}$$

$$C(\mathbf{x})p_n^c(\mathbf{x}) = \int_L [p_n^c H_n' - p_n'^c H_n] r_y dL_y \tag{25}$$

(where the superscript * has disappeared). These modified Kirchhoff–Helmholtz integral equations form an implicit expression for the surface pressure and its normal derivative. It can be used to determine the boundary values of p when p' is known and vice versa.[2]

II. COMPUTATIONAL ASPECTS OF FOURIER BEM

The solution of the Kirchhoff–Helmholtz equation for axisymmetric structures [Eq. (23)] can be obtained numerically by solving Eqs. (24) and (25) using standard boundary element procedures. The generator L of the axisymmetric body is discretized and the geometry and acoustic variables p and p' are assumed to vary according to isoparametric shape functions on the surface of the body. The discretization of the body involves only line elements. The evaluation of the integrals (4), (13), and (16) over the angular coordinate θ has to be sufficiently accurate and efficient, to use this method effectively.[1]

The possible singularities in the circumferential integrands of Eqs. (13) and (16) are not the only difficulties in their computation. The cosine function in the expressions for H_n and H'_n causes the total integrand to oscillate rapidly for high Fourier mode numbers n. Moreover, the $R(\mathbf{x},\mathbf{y})^{-1}$ function in these integrals causes a steep slope of the integrand near $\theta=0$ and $\theta=2\pi$, when the distance between $\mathbf{x}$ and $\mathbf{y}$ is relatively small. Therefore, special attention should be paid to the evaluation of these integrals.

Calculation of the circumferential integrals consumes a major portion of the total amount of computation time for the Fourier boundary element method. The integrals need to be computed often, and the calculation itself is computationally expensive. In general, the line integrals from Eqs. (24) or (25) need to be computed numerically for each Fourier harmonic n, for a number of observer points $\mathbf{x}$. This requires a value for H_n and H'_n and thus two circumferential integral evaluations on each integration point of the line integral, for each Fourier mode number n. In addition, applying fixed point numerical integration for these circumferential integrals, like Gaussian or Trapezium quadrature, requires a large number of integration points in circumferential direction to obtain sufficiently accurate results.[1] This is particularly true for high Fourier mode numbers n where the integrand evinces an oscillatory behavior. The long computation times resulting from the application of fixed point integration methods as proposed by Akyol,[1] Soenarko,[2] and Juhl[3] weaken the advantages of the Fourier BEM compared to the three-dimensional BEM. A solution to this problem will be presented in the next section.

A. Integral evaluation based on fast Fourier transform

For a sufficiently accurate and efficient evaluation of the integrals (13) and (16) for H_n and H'_n, respectively, a new method based on fast Fourier transform (FFT) was developed. The integrands in the expressions for H_n and H'_n consist of a reasonable smooth (but possibly singular) function multiplied by a cosine function. The integrand without the cosine function is an even periodic function around $\theta=0$ with a period equal to 2π. Given an even periodic function $f(x)$ with period $2L$, the Fourier coefficients F_n of this function are given by[5]

$$F_n = \frac{1}{L} \int_{-L}^{L} f(x)\cos\left(\frac{n\pi}{L}x\right)dx$$
$$= \frac{1}{L} \int_{0}^{2L} f(x)\cos\left(\frac{n\pi}{L}x\right)dx. \qquad (26)$$

With $x=\theta$ and $L=\pi$, this shows that Eqs. (13) and (16) are valid expressions of the Fourier coefficients H_n and H'_n of the complex even functions:

$$h(\theta) = \frac{e^{-ikR(\theta)}}{4R(\theta)} \qquad (27)$$

and

$$h'(\theta) = \frac{\partial}{\partial\nu}\left(\frac{e^{-ikR(\theta)}}{4R(\theta)}\right), \qquad (28)$$

respectively, which are in fact slightly modified forms of the three-dimensional free-space Green's function from Eq. (2) and its normal derivative. Hence, to compute the integrals in the expressions for H_n and H'_n we can use the nth Fourier coefficient of the complex functions h and h', respectively.

In numerical mathematics, the algorithm normally used for an efficient computation of the Fourier coefficients of a (complex) function is fast Fourier transform (FFT). FFT algorithms are optimized for speed while their accuracy is unaffected. Therefore, they are a good alternative for computing the integrals H_n and H'_n, but are generally more expensive than most fixed point numerical integration routines. However, a significant advantage of the proposed method is that by one FFT, the Fourier coefficients of many Fourier modes n are calculated, whereas the fixed point integration methods required an integral evaluation for every Fourier mode n.

An algorithm for the evaluation of the integrals in Eqs. (13) and (16) using FFT requires the following actions:

1. Determination of the number of samples n_{FFT} needed for computation of integrals H_n and H'_n with a desired accuracy.

2. Evaluation of the functions h [Eq. (27)] and h' [Eq. (28)] on n_{FFT} equidistant values of the parameter θ in the interval $[0, 2\pi]$.

3. Fast Fourier transformation of the n_{FFT} computed function values.

4. Selection of the nth terms of the calculated Fourier spectrums

which are numerical values for the integrals H_n and H'_n.

Numerical problems can occur when $\mathbf{x}$ coincides with $\mathbf{y}$ because the functions h and h' cannot be evaluated due to the R^{-1} singularity in Eqs. (27) and (28). This problem can be circumvented by taking surface integration points $\mathbf{y}$ that do not coincide with the observer point $\mathbf{x}$ when the line integrals (24) and (25) are computed. Gauss-log integration[6] should then be applied for the line integrals because the functions h and h' have logarithmic behavior near the singularity. Other methods proposed in literature to handle a singular integrand use a technique of subtracting and adding up the singular part of the integrand from the regular part, resulting in a regular and singular (surface) integral. Then, special (analytical) integration techniques[7,8] are used for the singular integral, and ordinary (Gauss–Legendre) integration techniques are used for the regular part. To apply this technique here, a special integration technique should be developed for the integral over a ring-shaped surface of the R^{-1} singularity. The integration region can be split up in an integral over the circumference and an integral over a generator segment of the vibrating body. The circumferential integral yields elliptic integrals of the first and second kind,[2] but an analytical solution for the integration of these elliptic integrals over the generator segment is not available. Therefore, this regularization method cannot be used here. It should however be noted that this is not a consequence of the pro-

posed FFT method but a general result for all FBEM implementations.

The application of the method of computing the circumferential integrals using FFT techniques is not limited to acoustic problems governed by the Helmholtz equation. In like manner, the new method is usable in non-acoustic Fourier BEM applications. The only condition to be fulfilled for the kernel is that it has to be a periodic function of θ.

B. Efficiency of the FFT method

Regarding efficiency, it is preferable that the number of evaluations of the functions $h(\theta)$ and $h'(\theta)$ can be chosen as low as possible, because this number is directly related to the cost of the FFT algorithm in particular and the cost of the total method in general. The number of required evaluations of these functions is determined by the desired accuracy of the FFT process. Signal leakage and aliasing in the Fourier transform process should be taken into account. This means that precisely an integer number of periods of the periodic functions $h(\theta)$ and $h'(\theta)$ should be sampled, and that the sampling frequency should at least be twice the highest frequency present in the functions. The first requirement is easily met because we exactly know the period of the functions $h(\theta)$ and $h'(\theta)$ which is 2π for axisymmetric structures. To satisfy the requirements for aliasing, however, the frequency content of functions $h(\theta)$ and $h'(\theta)$ needs to be predicted because it is not known beforehand. This is the topic of the remaining part of this section. For numerical efficiency it is desirable that the number of Fourier transform points can be written as 2^f with f as a positive integer number.

A closer look at the functions h and h' that are Fourier transformed is illustrative to establish a reasonable expression for the minimum number of Fourier transform points n_{FFT} required for each integral evaluation. A representative picture of the function h is plotted in Fig. 2.

The steepness of the curve close to $\theta=0$ and $\theta=2\pi$ is determined by the ratio of the minimum and maximum R (distance between the points $\mathbf{x}$ and $\mathbf{y}$), because of the factor R^{-1} in Eq. (27) for h. The ratio R_{max}/R_{min} can serve as a dimensionless scale factor for the problem's geometry. The rapid change in steepness of the curve for large values of this ratio causes nonzero high-mode number components in the Fourier transform of the function h. Therefore a sufficiently high number of Fourier points has to be used. Hence, a criterion for the minimum number of Fourier points needed should be a function of a steepness parameter:

$$c_s = \frac{R_{max}}{R_{min}}. \tag{29}$$

The oscillations in the curves for the real and imaginary part are caused by the term $e^{-ikR(\theta)} = \cos(kR) - i\sin(kR)$ for large numbers of k and/or a large difference between the minimum and maximum value for R. Thus the criterion for the minimum number of Fourier points should also be a function of an oscillation parameter:

$$c_0 = k(R_{max} - R_{min}). \tag{30}$$
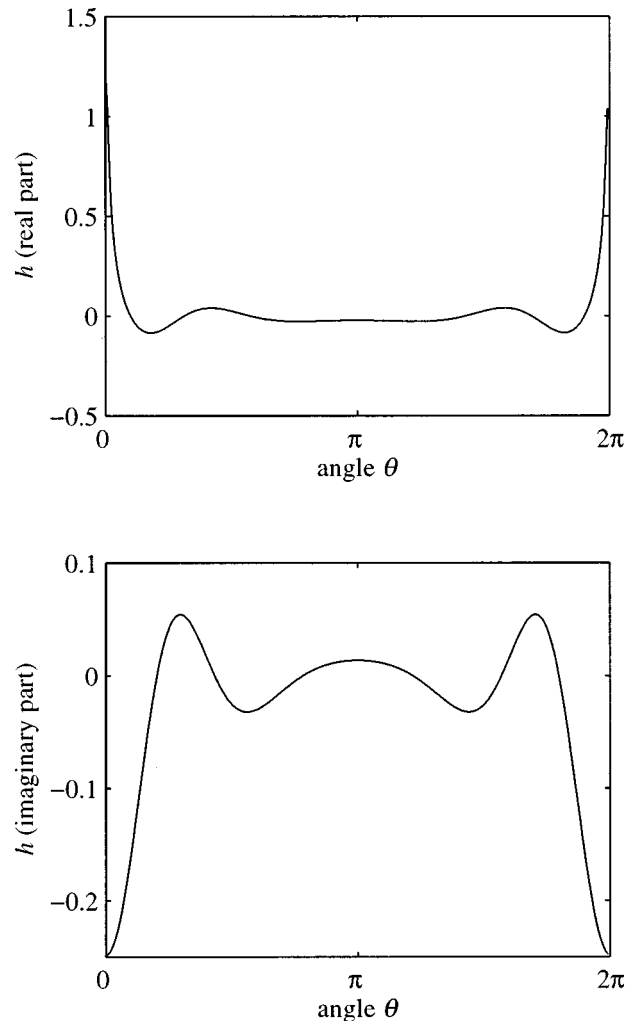
FIG. 2. Function h for $X=(5,0,0)$, $Y=(5,0,0.2)$ and $k=1$.

The same considerations can be made for the Fourier transform of h'_n, leading to the same parameters c_s and c_0.

An expression for the minimum number of Fourier points is dependent on the characteristics of the integrands h_n and h'_n which can be described by the parameters c_s and c_0. So an expression for n_{FFT} can be expressed as a function of those parameters:

$$n_{FFT} = n_{FFT}(c_s, c_0) = n_{FFT}(R_{max}, R_{min}, k) = n_{FFT}(\mathbf{x}, \mathbf{y}, k). \tag{31}$$

For an efficient application of the FFT method, an expression for n_{FFT} can be developed for a desired accuracy. The expression that can be derived is generally applicable for efficiently computing integrals in Eqs. (13) and (16) with FFT integration. For efficiency, it is also important to obtain a relatively simple expression n_{FFT}, because it must be used for each FFT based integral evaluation separately. Fortunately, a simple expression can be derived for practical application, as is illustrated in Sec. III.

C. Discretization process

The boundary element method is applied for the discretization of the modified Kirchhoff–Helmholtz integral equations (24) and (25). Assume that the generator L of body B

can be discretized with n_e line elements and that each line element i_e has n_n nodes. The total number of nodes is denoted as n_{nd}. Thus the coordinates r and z can be expressed in terms of the coordinates r_{i_n} and z_{i_n} of element node i_n using a piecewise polynomial approximation:

$$r(\xi) = \sum_{i_n=1}^{n_n} N_{i_n}^{n_n}(\xi) r_{i_n}, \tag{32}$$

$$z(\xi) = \sum_{i_n=1}^{n_n} N_{i_n}^{n_n}(\xi) z_{i_n}, \tag{33}$$

where $N_{i_n}^{n_n}(\xi)$ are the (n_n-1)th order isoparametric shape functions, ξ is the local element coordinate and i_n is the local node number. The boundary variables p_0, p_0', p_n^s, p_n^c, $p_n'^s$, and $p_n'^c$ are also approximated using the same isoparametric shape functions as for the coordinates, thus on element i_e,

$$\psi_{i_e}(\xi) = \sum_{i_n=1}^{n_n} N_{i_n}^{n_n}(\xi) \psi_{i_e i_n}, \tag{34}$$

where any of the boundary variables can be substituted for $\psi_{i_e}(\xi)$, and $\psi_{i_e i_n}$ is the value of the corresponding variable on local node i_n of element i_e. Using this approximation in, for instance, Eq. (24), yields

$$C(\mathbf{x}) p_n^s(\mathbf{x}) = \sum_{i_e=1}^{n_e} \left\{ \sum_{i_n=1}^{n_n} p_{n i_e i_n}^s \int_{-1}^1 N_{i_n}^{n_n}(\xi) H_n'(\xi) r(\xi) J_{i_e}(\xi) d\xi \right.$$
$$\left. - \sum_{i_n=1}^{n_n} p_{n i_e i_n}'^s \int_{-1}^1 N_{i_n}^{n_n}(\xi) H_n(\xi) r(\xi) J_{i_e}(\xi) d\xi \right\}, \tag{35}$$

where $p_{n i_e i_n}^s$ is the value of p_n^s at local node i_n of element i_e and $J_{i_e}(\xi)$ the Jacobian of the transformation given by Eqs. (32) and (33), for element i_e:

$$J_{i_e}(\xi) = \left[\left(\frac{dr}{d\xi} \right)^2 + \left(\frac{dz}{d\xi} \right)^2 \right]^{1/2}. \tag{36}$$

Expressions similar to Eq. (35) can be obtained for the other boundary values using Eq. (25).

For the solution process, a collocation scheme is applied. The observer points $\mathbf{x}$ on the boundary are chosen successively to coincide with each global node i_{nd} and $\mathbf{y}$ is the (surface) point of integration, now explicitly a function of ξ through Eqs. (32) and (33). This collocation method results in a set of n_{nd} linear algebraic equations in terms of the unknown p_n, when p_n' is given on each node, and vice versa. The resulting equations may be written in the following matrix form:

$$\mathbf{A_n} p_n^s = \mathbf{B_n} p_n'^s, \tag{37}$$

where p_n^s and $p_n'^s$ are the column vectors containing the n_{nd} nodal values of p_n^s and p_n^0, respectively. $\mathbf{A_n}$ and $\mathbf{B_n}$ are square matrices with the various integrals as in Eq. (35) as their elements. For the cosine terms of the Fourier series, a similar matrix equation can be derived:

$$\mathbf{A_n} p_n^c = \mathbf{B_n} p_n'^c, \tag{38}$$

where p_n^c and $p_n'^c$ are the column vectors containing the n_{nd} nodal values of p_n^c and $p_n'^c$. Thus, two matrix equations result, relating the terms of the Fourier expansion of the unknown variables to the Fourier coefficients of the boundary conditions. The matrix equations that describe the acoustic radiation have to be formed for every Fourier mode number n that is present in the Fourier expansion of the boundary conditions.

When the solution for all boundary values is computed, the value of the Fourier terms of the pressure and its derivative for any surface or exterior point can easily be obtained by applying an equation similar to Eq. (35) for p_n^{s*} and p_n^{c*}, substituting the calculated values for $p_{n i_e i_n}^s$, $p_{n i_e i_n}^c$, $p_{n i_e i_n}'^s$, and $p_{n i_e i_n}'^c$. The resulting equation is an explicit relationship between the acoustic variables on the surface of the radiating body and the acoustic variables at any other position in the acoustic medium.

III. APPLICATION OF THE FFT METHOD

The FFT based integration method was implemented in the in-house acoustic Fourier BEM code BARD to assess its efficiency compared to a Gaussian quadrature based method that was previously proposed. For the FFT method, an expression for the minimum number of Fourier points n_{FFT} was developed. For a large number of parameters values c_s in the range $[1,2000]$ and c_0 in the range $[0,100]$, the integrals (13) and (16) were computed with the proposed method until convergence was achieved. Each parameter pair c_s and c_0 has a specific minimum number of Fourier points n_{FFT} for the integrals to converge where the relative error in the computed value for the integrals did not exceed 10^{-3}. These pairs of c_s, c_0 and related n_{FFT} were then used in a curve-fitting procedure to obtain the relationship

$$n_{FFT}(c_s, c_0) = 14 \left(c_s + \frac{c_0}{2\pi} \right)^{0.9}.$$

In the BEM code the expression for n_{FFT} was implemented and its value is computed for each circumferential integral evaluation seperately based on values for c_s and c_0 for that integral. Its value is rounded to the nearest subsequent power of 2, to enable the use of a fast radix-2 FFT algorithm. An efficient integration algorithm based on Gaussian quadrature was also implemented such that the maximum relative error with that method was 10^{-3}.

As a check for the accuracy of the developed FBEM code BARD, the response of an oscillating sphere with radius $a = 1$ was computed. To model the sphere, a mesh with 10 quadratic line elements was used and the response was calculated for a dimensionless wavenumber range $ka = 0,...,5$. The relative errors in the computed surface pressure obtained with the FBEM compared with the analytically known solution for this problem[9,10] were smaller than 1%.

To assess the efficiency of the FFT based integration method, the acoustic responses of two model vibrating bodies, a sphere and cylinder, were computed. The sphere and
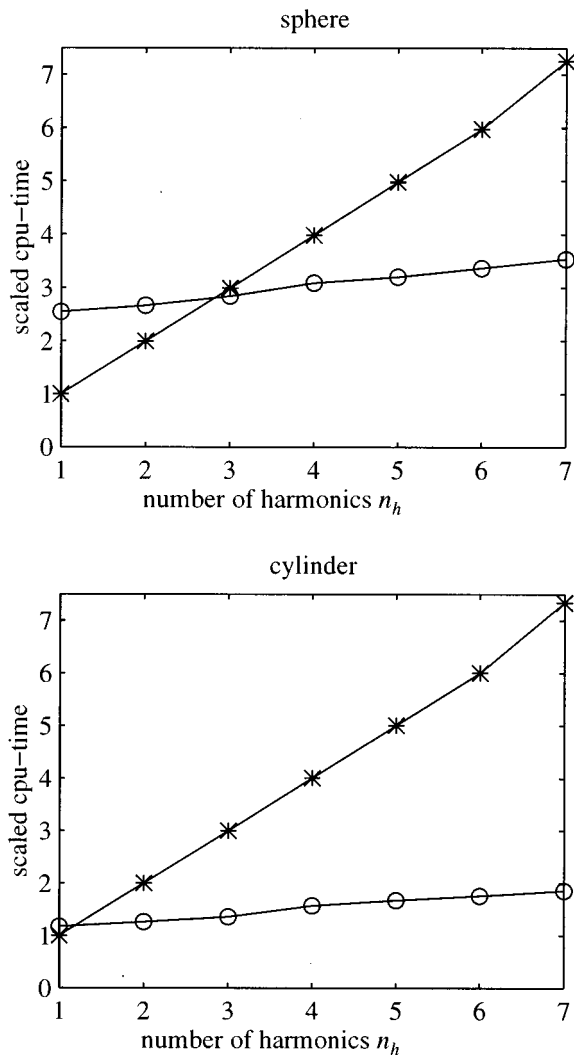
FIG. 3. Scaled computation time for spherical and cylindrical radiation for Gauss (*) and FFT (○) based quadrature.

cylinder were discretized using quadratic elements. The normal velocity boundary condition for both problems was defined as

$$u_\nu(\mathbf{y}) = \frac{1}{-i\rho\omega} p'(\mathbf{y}) = \sum_{n=1}^{n_h} \cos(n\theta),$$

where the number of harmonics n_h ranged from 1 to 7 to assess the usability of the methods for constructions with more complicated boundary conditions in circumferential direction. The acoustic pressure response was computed with both the FFT based and Gauss based integration methods. The differences between the results obtained with the FFT method and the Gauss method were negligible, indicating that the circumferential integrals were computed with the same accuracy. The efficiency of the new method can be assessed by making a comparison between cpu time needed for the radiation computation for both problems as a function of the number of harmonics n_h in the boundary conditions as depicted in Fig. 3.

The cpu times were scaled by dividing the actual computation time by the time needed to compute the response of the first harmonic using the Gaussian quadrature method. For both methods and for both problems, the cpu time needed to compute the acoustic response increased linearly with each extra harmonic in the boundary conditions. However, the cpu time for the Gaussian quadrature method increased more per extra harmonic than for the FFT method. This can be explained by the most important difference between the methods: when the Gauss method is used, the circumferential integrals are computed for each Fourier mode separately, but when the FFT method is used, these integrals are computed for all modes simultaneously. In other words, the characteristic total cpu time for the radiation computation with the Gauss method is the computation time for one harmonic multiplied by the total number of harmonics in the boundary conditions, while for the FFT method, the computation time for evaluating *one* harmonic is typical of the total cpu time for all harmonics. For one or two harmonics in the boundary conditions for the sphere radiation, and for one harmonic in the cylinder boundary conditions, the calculations with Gaussian quadrature were faster than those with the FFT method (Fig. 3). However, for *all* other problem configurations that were tested, using the FFT method was clearly advantageous. In conclusion, for more than one harmonic in the boundary conditions, the FFT method is generally more efficient than the Gaussian quadrature method, while obtaining the same accuracy.

IV. CONCLUSIONS

A new method was presented to handle the integrals over the angle of revolution arising in the Fourier boundary element method (FBEM) for the calculation of the acoustic response of axisymmetric bodies with arbitrary boundary conditions. Previously proposed methods to handle these integrals required a separate (expensive) integral evaluation for each Fourier term (circumferential harmonic) that is present in the boundary conditions. With the integration method proposed here, these integrals are computed simultaneously for all harmonics. The method efficiently uses fast Fourier transform (FFT) algorithms to obtain numerical values for these integrals.

Simulations were performed comparing the efficiency of the new FFT integration method with that of a method based on Gaussian quadrature. They showed that the FFT method is generally more efficient, especially when more than one circumferential harmonic is present in the boundary conditions.

It is expected that the utility of the new integration method presented here is not limited to acoustic problems only. The method should be usable whenever the integrand of the circumferential integral is periodic and (fast) Fourier transformable.

[1] T. P. Akyol, ''Schallabstrahlung von Rotationskörpern,'' Acustica **61**, 200–212 (1986).

[2] B. Soenarko, ''A boundary element formulation for radiation of acoustic waves from axisymmetric bodies with arbitrary boundary conditions,'' J. Acoust. Soc. Am. **93**, 631–639 (1993).

[3] P. Juhl, ''An axisymmetric integral equation formulation for free space non-axisymmetric radiation and scattering of a known incident wave,'' J. Sound Vib. **163**, 397–406 (1993).

[4] R. D. Ciskowski and C. A. Brebbia, Eds., *Boundary Element Methods in Acoustics* (Elsevier Applied Science, London, 1991).

[5] E. Kreyszig, *Advanced Engineering Mathematics* (Wiley, New York, 1993), 7th ed.

[6] M. H. Lean and A. Wexler, ''Accurate numerical integration of singular boundary element kernels over boundaries with curvature,'' Int. J. Numer. Methods Eng. **21**, 211–228 (1985).

[7] M. Guiggiani and P. Casalini, ''Direct computation of cauchy principal value integrals in advanced boundary elements,'' Int. J. Numer. Methods Eng. **24**, 1711–1720 (1987).

[8] H. Liao and Z. Xu, ''A method for direct evaluation of singular integral in direct boundary element method,'' Int. J. Numer. Methods Eng. **35**, 1473–1485 (1992).

[9] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton U.P., Princeton, NJ, 1968).

[10] E. Skudrzyk, *The Foundations of Acoustics* (Springer-Verlag, Vienna, 1971).

General theory for the spectral evolution of nonlinear Rayleigh waves

E. Yu. Knight, M. F. Hamilton, Yu. A. Il'inskii, and E. A. Zabolotskaya
Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712-1063

The spectral theory developed by Zabolotskaya [J. Acoust. Soc. Am. **91**, 2569–2575 (1992)] for nonlinear Rayleigh waves in isotropic solids is generalized to include fields with nonperiodic waveforms and nonplanar spatial dependencies. Model equations for cylindrical waves and diffracting beams that have been used in previous work, and which were obtained in an *ad hoc* manner, are derived here rigorously from first principles. The present theory for plane waves, which is derived using Hamiltonian formalism, is shown to be equivalent to the theory obtained by Parker [Int. J. Eng. Sci. **26**, 59–75 (1988)], whose derivation is based on a different approach. © *1997 Acoustical Society of America.* [S0001-4966(97)04508-6]

PACS numbers: 43.25.Fe [MAB]

INTRODUCTION

One purpose of the present article is to generalize Zabolotskaya's derivation[1] of model equations for periodic, plane, nonlinear Rayleigh waves in isotropic solids to include waveforms that exhibit an arbitrary dependence on time. The derivation presented here is also more rigorous in several respects. As discussed below, model equations for plane nonlinear Rayleigh waves have been developed by a number of other researchers. However, the derivation in Ref. 1 is the first to be based on Hamiltonian formalism, which is substantially different from the approaches taken by others. The form of the model equation that we obtain, although equivalent to a result obtained previously by Parker[2] (which we prove in Appendix A), in our opinion lends itself more readily to physical interpretation (Parker's nonlinear term is quite different in appearance), and the approximations made in the present derivation are more easily identified.

A second purpose is to derive from first principles the model equations that account for geometrical spreading and diffraction, as well as arbitrary time dependence, in nonlinear Rayleigh waves. A rigorous derivation of a model equation for nonplanar, nonlinear Rayleigh waves has not appeared previously in the literature. Although effects of cylindrical spreading and diffraction have been included in earlier nonlinear model equations, the terms corresponding to these effects were introduced *ad hoc*. As in our derivation for plane Rayleigh waves, we use Hamiltonian formalism to develop equations for nonplanar Rayleigh waves, but the latter derivation requires a slightly different approach. The derivation reveals very clearly the restrictions on use of existing model equations for nonplanar nonlinear Rayleigh waves.

Before proceeding with our derivations of the model equations, we begin with a brief review of theoretical models developed by others.

The linear solution for a surface acoustic wave (SAW) propagating in an isotropic solid along a traction-free interface dates back to the work of Lord Rayleigh[3] in 1885. The first nonlinear studies were performed at a much later time. Theoretical predictions made by Viktorov[4] in 1964 and experimental results reported in the following year by Rischbieter[5,6] were the first attempts to investigate harmonic generation in surface waves. During the last three decades, interest in nonlinear Rayleigh waves was motivated primarily by the use of SAW devices for nonlinear signal processing applications (see, for example, Ref. 7). More recently, Kolomenskii and Maznev[8] demonstrated that lasers can generate intense surface wave pulses having predicted shock formation distances on the order of centimeters.[9]

Early theories for the nonlinear propagation of surface waves were mainly phenomenological. The models developed by Lean and Tseng[10] and by Adler *et al.*[11] contain nonlinear coupling coefficients that must be determined by comparison with experimental results, instead of being related directly to properties of the material. Several theoretical models were proposed where the coupling coefficients are related to second- and third-order elastic constants by use of the full second-order theory of elasticity. The difficulty of the problem, however, prevented the early researchers from satisfying both the nonlinear equations of motion and the nonlinear boundary condition to the same order of approximation. As pointed out later by Kalyanasundaram,[12] asymptotic solutions obtained by Løpen[13] and Ljamov *et al.*[14] do not satisfy the stress-free boundary condition at the chosen order of approximation, and the solutions obtained by Tiersten *et al.*[15] and Vella *et al.*[16] satisfy the boundary condition but not the governing equations.

Not until the 1980's was a rigorous approach finally undertaken, when Kalyanasundaram[12] removed the limitations of previous theories arising from neglecting the singular nature of boundary matrices at each order of approximation. Consistent coupled equations were derived for second harmonic generation using the method of multiple scales. The obtained theory was still limited because the asymptotic solution was truncated at an early stage, with only the fundamental and second harmonic components taken into account. This restriction was removed by Kalyanasundaram *et al.*[17] in a subsequent paper, where an infinite set of coupled equations was derived for the spectral amplitudes in a periodic plane wave. Perturbation theory was used to obtain a set of linear inhomogeneous differential equations with

inhomogeneous boundary conditions. Substitution of the solution of the equations of motion into the boundary condition led to a system of linear inhomogeneous algebraic equations at each order of the perturbation expansion. The coefficient matrices of the obtained algebraic equations are singular because surface waves are nondispersive. Thus the requirement that there exist a solution of these algebraic equations imposes appropriate solvability conditions on the inhomogeneous terms in the equations. Use of the solvability, or consistency, conditions led to the set of coupled spectral equations. The derived model equations were used to investigate effects such as second harmonic generation, sum and difference frequency generation (all in the quasilinear approximation), and waveform distortion of an initially sinusoidal signal. The existence of a steady-state periodic solution was investigated by Kalyanasundaram and Anand,[18] and the analysis was also extended by Kalyanasundaram[19] to the problem of interaction of counterpropagating waves, which are exploited in various nonlinear SAW device applications.

Lardner[20] extended the theory developed by Kalyanasundaram to include transient waveforms. He also introduced a slow scale in the depth coordinate, which allowed him to eliminate the secular terms proportional to the transverse coordinate, and a uniformly valid solution was thus obtained. Effects of thermal conductivity and internal damping (with a viscous term of the Kelvin–Voigt type) were included by Lardner in a subsequent analysis.[21] Harmonic generation, parametric mixing, and waveform distortion were investigated analytically and numerically within the framework of the developed theory,[20–22] although the material constants that were used do not correspond to any specific real material. A similar approach was undertaken by David,[23] who introduced a slow scale for the transverse coordinate, which resulted in a uniformly valid solution for the interior of a solid. Harmonic generation and waveform distortion were investigated numerically by David for the case of fused quartz.

The existence of nonlinear Rayleigh waves of permanent form in an isotropic solid was investigated by Parker and Talbot.[24] Equations were obtained specifically for stationary surface waves. A general theory for plane surface wave propagation in an isotropic elastic solid, as well as in media possessing anisotropy, was developed by Parker.[2] His analysis was based on a dynamical approach, perturbation theory was used, and coupled equations were obtained from the compatibility condition for the slow modulation of the Fourier transform of the signal due to nonlinearity. Explicit equations were presented only for isotropic materials. Predictions obtained from the model were used to investigate the distortion of an initially sinusoidal signal in the harmonic material introduced by John.[25] The theory was later extended by Kalyanasundaram et al.[26] to model the propagation of nonlinear surface wave beams.

The theories developed by Kalyanasundaram, Lardner, and Parker are based on the idea of obtaining the coupled spectral equations as the compatibility condition that emerges in the solution of a nonlinear boundary value problem by perturbation methods. An alternative approach was suggested by Shui and Solodov.[27] Nonlinear Rayleigh and Stoneley wave propagation was treated as a degenerate case of bulk acoustic wave reflection at the interface. The coupled equations were obtained in the quasilinear approximation.

As noted above, Zabolotskaya's methodology[1] differs from earlier analytical techniques, all of which are based on a dynamical approach to the problem, and which result in a nonlinear equation with a nonlinear boundary condition. In contrast, the theory developed in Ref. 1 is based on Hamiltonian mechanics, with a set of spatial Fourier amplitudes used for generalized coordinates, and with the stress-free boundary condition built into the formalism. The original governing equations were derived for periodic, plane, progressive surface waves propagating in an isotropic elastic solid. Equations for the spectral amplitudes are obtained directly from the Hamiltonian canonical equations. The theory was used in Ref. 1 to investigate distortion and shock formation in plane and cylindrical Rayleigh waves that propagate in a real material, in particular, steel 60 C2H2A.

Several theoretical investigations[28–32] of nonlinear Rayleigh waves have been performed on the basis of Zabolotskaya's model equations.[1] The existence of stationary waves was studied by Hamilton et al.[28] in response to the article by Parker and Talbot.[24] It was concluded that Rayleigh waves of permanent form exist only if artificial constraints are imposed on the frequency spectrum of the signal. Harmonic generation, acoustic saturation, and finite amplitude losses in plane surface waves were investigated analytically and numerically by Shull et al.[29] An analytical prediction of the shock formation distance was proposed on the basis of comparison of the quasilinear solution for second harmonic generation with the corresponding solution for a finite amplitude sound wave in an ideal fluid. The propagation of cylindrical waves[29] and diffracting beams[30] was also studied, where the spreading and diffraction terms in the model equations were introduced ad hoc. Analytical studies of the corresponding time-domain evolution equation for the horizontal velocity component at the interface were also performed. The kernel of the nonlinear integral operator in the evolution equation was investigated and separated into two parts, one associated with local and the other with nonlocal nonlinearity.[31] The obtained nonlinear operator corresponding to the local part of the kernel has exactly the same functional dependence as for finite amplitude sound in fluids. A different form of the evolution equation, expressed in terms of the horizontal displacement component, was also derived.[32]

In the present paper, Zabolotskaya's derivation[1] of the basic spectral equations for nonlinear Rayleigh wave propagation in isotropic solids is generalized to include aperiodic signals, and to account for spreading losses and diffraction effects. Plane waves having arbitrary time dependence are considered in Sec. I. In Sec. II, equations for nonplanar waves are derived rigorously, and simplified models for cylindrical waves and diffracting beams are obtained. In Appendix A it is shown that the present theory is equivalent to one developed earlier by Parker,[2] who used a substantially different analytical approach. In Appendix B, discrete forms of the model equations derived in the present article are pro-

vided. The discrete forms are more amenable to numerical solution.

I. PLANE WAVES

Here we consider a plane wave, not necessarily periodic, that propagates along an infinite plane surface of a lossless, isotropic elastic solid. The coordinate system is such that the surface coincides with the (x,y) plane, and the solid occupies the half-space $z \leqslant 0$. The plane waves propagate along the x axis.

We assume that the solution of the nonlinear wave propagation problem is very close to the linear solution, in particular, that the depth dependence of each frequency component is the same as in the linear solution. The physical explanation for this assumption is revealed by considering the exact solution of the nonlinear problem as an expansion over all possible elastic wave modes, bulk (compressional and shear) modes as well as surface modes. Although there is no coupling between these modes in the linear approximation, this is not the case for a nonlinear system, in which energy transfer between surface waves and bulk waves is possible. The coupling results in a different mode structure for the solution, which can no longer be represented as a superposition of surface waves alone. Consequently, the depth dependence of each frequency component differs from that in the linear solution. However, interactions between bulk and surface waves are nonresonant, and therefore inefficient, because of the differences in propagation speeds. These interactions are ignored because the error introduced by this approximation is of the same order as other terms that are ignored. As a result of this observation, the depth dependence of each harmonic mode can be taken as for a small-signal surface wave.

A. Basic equations

For a plane surface wave (not necessarily progressive) that propagates along the x axis, all field variables are independent of the y coordinate, and the particle displacement vector has only x and z components. The horizontal and vertical displacement components, u_x and u_z, respectively, are expressed through Fourier integrals as follows:

$$u_x(x,z,t) = \int_{-\infty}^{\infty} i \frac{k}{|k|} a_k(t) h_x(z,k) e^{ikx} \frac{dk}{2\pi}, \qquad (1)$$

$$u_z(x,z,t) = \int_{-\infty}^{\infty} a_k(t) h_z(z,k) e^{ikx} \frac{dk}{2\pi}. \qquad (2)$$

Here, $a_k(t)$ is the spectral density of the mode with wave number k, and the depth-dependent functions are given by

$$h_x(z,k) = \xi_t e^{|k|\xi_t z} + \eta e^{|k|\xi_l z}, \qquad (3)$$

$$h_z(z,k) = e^{|k|\xi_t z} + \eta \xi_l e^{|k|\xi_l z}, \qquad (4)$$

where $\xi_t = (1-\xi^2)^{1/2}$, $\xi_l = (1-\xi^2 c_t^2/c_l^2)^{1/2}$, $\eta = -2(1-\xi^2)^{1/2}/(2-\xi^2)$, $\xi = c_R/c_t$, c_l and c_t are the phase speeds of com-

pression and shear bulk waves, respectively, and c_R is the small-signal Rayleigh wave speed. Note that the displacement field is defined in terms of its Fourier transform in space, and the variation of a_k with time is the unknown quantity to be determined. When the time dependence of a_k is $e^{\pm ikc_R t}$, Eqs. (1) and (2) are the Rayleigh wave solutions in the linear approximation.

Following the technique developed in Ref. 1, we shall employ Hamiltonian mechanics to derive a set of coupled equations for the evolution of the frequency spectrum of a nonlinear Rayleigh wave. Therefore, the first step is to calculate the kinetic and elastic energies of the system in terms of the spectral functions $a_k(t)$. Strains in the material are assumed to be small and characterized by the parameter $\epsilon = u_0/L \ll 1$, where u_0 is the magnitude of a characteristic particle displacement, and L is a length scale corresponding to a characteristic wavelength. Quadratic nonlinearity of the field variables shall be taken into account, and therefore terms up to order ϵ^3 are retained in the expression for the elastic energy.

The kinetic energy density (i.e., per unit length in the y direction) T for the signal given by Eqs. (1) and (2) is defined by

$$T = \int_{-\infty}^{\infty} \int_{-\infty}^{0} \frac{\rho}{2} (\dot{u}_x^2 + \dot{u}_z^2) dx \ dz$$

$$= \frac{\rho \zeta}{4} \int_{-\infty}^{\infty} \dot{a}_k \dot{a}_{-k} |k|^{-1} \frac{dk}{2\pi}, \qquad (5)$$

where ρ is the ambient density, dots indicate time derivatives, and

$$\zeta = \xi_t + \frac{1}{\xi_t} + \eta^2 \xi_l + \frac{\eta^2}{\xi_l} + 4\eta. \qquad (6)$$

Calculation of the quadratic part of the potential energy density V yields[1]

$$V = \int_{-\infty}^{\infty} \int_{-\infty}^{0} \mathscr{E}_2 \ dx \ dz = \frac{\rho c_R^2 \zeta}{4} \int_{-\infty}^{\infty} a_k a_{-k} |k| \frac{dk}{2\pi}, \quad (7)$$

where $\mathscr{E}_2$ is the quadratic part of the elastic energy density (per unit volume) in coordinate space. The expression for the cubic part of the potential energy density W, which accounts for nonlinear wave interactions, becomes

$$W = \int_{-\infty}^{\infty} \int_{-\infty}^{0} \mathscr{E}_3 \ dx \ dz$$

$$= 2\pi\mu \int \int \int_{-\infty}^{\infty} w(k_1,k_2,k_3) a_{k_1} a_{k_2} a_{k_3}$$

$$\times \delta(k_1 + k_2 + k_3) \frac{dk_1 \ dk_2 \ dk_3}{(2\pi)^3}, \qquad (8)$$

where $\mathscr{E}_3$ is the cubic part of the elastic energy density, μ is the shear modulus, δ is the Dirac delta function, and

$$w(k_1,k_2,k_3)=-\frac{|k_1k_2k_3|}{3}\left(\frac{\alpha'}{|k_1|\xi_t+|k_2|\xi_t+|k_3|\xi_t}+\frac{\alpha'}{|k_1|\xi_t+|k_2|\xi_t+|k_3|\xi_t}+\frac{\alpha'}{|k_1|\xi_t+|k_2|\xi_t+|k_3|\xi_t}\right.$$

$$\left.+\frac{\beta'}{|k_1|\xi_t+|k_2|\xi_l+|k_3|\xi_l}+\frac{\beta'}{|k_1|\xi_l+|k_2|\xi_t+|k_3|\xi_l}+\frac{\beta'}{|k_1|\xi_l+|k_2|\xi_l+|k_3|\xi_t}+\frac{3\gamma'}{|k_1|\xi_l+|k_2|\xi_l+|k_3|\xi_l}\right). \tag{9}$$

Only the symmetrical (with respect to permutation of k_1, k_2, and k_3) part of the expression for $w(k_1,k_2,k_3)$ is presented here, because the antisymmetrical part does not contribute to the integral for W in Eq. (8), and it can therefore be discarded. Equation (9) is related to the expression for $w_{k_1k_2k_3}$ given by Eq. (19) in Ref. 1 as

$$w(k_1,k_2,k_3)=\tfrac{1}{3}(w_{k_1k_2k_3}+w_{k_2k_3k_1}+w_{k_3k_1k_2}). \tag{10}$$

It can be seen from Eq. (8) that $w(k_1,k_2,k_3)$ is associated with the strength of the nonlinear interaction of spectral components with wave numbers k_1, k_2, and k_3. Explicit expressions for $\mathscr{E}_2$ and $\mathscr{E}_3$ follow directly from Eq. (11) in Ref. 1. The coefficients α', β', and γ' are also the same as in Ref. 1; their definitions in terms of both Landau and Murnaghan notation are provided in Appendix A of the present article.

The Hamiltonian of the system, up to cubic order in the field variables, is

$$H=H_0+W, \tag{11}$$

where $H_0=T+V$ is the Hamiltonian in the linear approximation (i.e., at quadratic order in the field variables).

B. Hamiltonian formalism

We now identify appropriate canonical variables for the system. The spectrum given by the set of functions a_k in Eqs. (1) and (2) is continuous rather than discrete (a discrete spectrum is assumed in Ref. 1). A continuum of generalized coordinates is therefore required. To describe a system in an infinite-dimensional phase space rather than in one that has finite dimensions, functional derivatives rather than partial derivatives must be used in the Hamiltonian canonical equations.

A natural choice of generalized coordinates is the infinite set of functions $\{a_k\}$, in which case the corresponding canonical momenta are defined by

$$p_k=\frac{\delta T}{\delta \dot{a}_k}=\frac{\rho\zeta}{4\pi|k|}\dot{a}_{-k}. \tag{12}$$

The set $\{a_k(t),p_k(t)\}$ defines the infinite-dimensional phase space describing the displacement field and its evolution with time. Equation (5) for the kinetic energy density, when expressed in terms of p_k, becomes

$$T=\frac{(2\pi)^2}{\rho\zeta}\int_{-\infty}^{\infty}p_kp_{-k}|k|\frac{dk}{2\pi}. \tag{13}$$

With the Hamiltonian expressed in terms of a_k and p_k, Hamilton's canonical equations of motion take the form

$$\dot{a}_k=\frac{\delta H}{\delta p_k}, \quad \dot{p}_k=-\frac{\delta H}{\delta a_k}. \tag{14}$$

Coupled equations for a_k and p_k can be obtained by calculating the derivatives in Eqs. (14) explicitly, but we find it more convenient to introduce a new set of variables $\{b_k,b^*_{-k}\}$ defined as follows:[33]

$$a_k(t)=b_k(t)e^{-i|k|c_Rt}+b^*_{-k}(t)e^{i|k|c_Rt}, \tag{15}$$

$$\dot{a}_k(t)=\frac{4\pi|k|}{\rho\zeta}p_{-k}$$
$$=-i|k|c_Rb_k(t)e^{-i|k|c_Rt}+i|k|c_Rb^*_{-k}(t)e^{i|k|c_Rt}. \tag{16}$$

Notice that b_k and b^*_{-k} are chosen such that the relation $a_{-k}=a^*_k$ holds, which follows from the requirement that the displacement components be real. Using Eqs. (15) and (16), we can express b_k through a_k and p_k as follows:

$$b_k=\left(\frac{a_k}{2}+\frac{2\pi i}{\rho c_R\zeta}p_{-k}\right)e^{i|k|c_Rt}. \tag{17}$$

Differentiation of Eq. (17) with respect to time gives

$$\dot{b}_k=i|k|c_Rb_k+\left(\frac{\dot{a}_k}{2}+\frac{2\pi i}{\rho c_R\zeta}\dot{p}_{-k}\right)e^{i|k|c_Rt}. \tag{18}$$

After Eqs. (14) are substituted into the right-hand side of Eq. (18), the latter takes the form

$$\dot{b}_k=i|k|c_Rb_k+\left(\frac{1}{2}\frac{\delta H}{\delta p_k}-\frac{2\pi i}{\rho c_R\zeta}\frac{\delta H}{\delta a_{-k}}\right)e^{i|k|c_Rt}. \tag{19}$$

Changing the variables in Eq. (19) to $\{b_k,b^*_{-k}\}$, we obtain

$$\dot{b}_k=i|k|c_Rb_k+\frac{1}{2}e^{i|k|c_Rt}\left(\frac{\delta H}{\delta b_{-k}}\frac{\partial b_{-k}}{\partial p_k}+\frac{\delta H}{\delta b^*_k}\frac{\partial b^*_k}{\partial p_k}\right)$$
$$-\frac{2\pi i}{\rho c_R\zeta}e^{i|k|c_Rt}\left(\frac{\delta H}{\delta b_{-k}}\frac{\partial b_{-k}}{\partial a_{-k}}+\frac{\delta H}{\delta b^*_k}\frac{\partial b^*_k}{\partial a_{-k}}\right). \tag{20}$$

Calculation of the partial derivatives of b_{-k} and b^*_k with respect to a_k and p_k, by use of Eq. (17) and its complex conjugate, reduces Eq. (20) to

$$\dot{b}_k=i|k|c_Rb_k-\frac{2\pi i}{\rho c_R\zeta}\frac{\delta H}{\delta b^*_k}. \tag{21}$$

When the energy densities T and V, Eqs. (5) and (7), respectively, are expressed in terms of b_k and b^*_{-k}, the Hamiltonian in Eq. (11) takes the form

$$H=\rho c_R^2\zeta\int_{-\infty}^{\infty}b_kb^*_k|k|\frac{dk}{2\pi}+W. \tag{22}$$

Substitution of Eq. (22) into Eq. (21) yields

$$\dot{b}_k=-\frac{2\pi i}{\rho c_R\zeta}\frac{\delta W}{\delta b^*_k}. \tag{23}$$

Equation (23) is correct to quadratic order in the sense that no approximation was made so far, apart from the assumption that surface waves do not interact with bulk waves.

We now express the right-hand side of Eq. (23) explicitly in terms of displacements. Writing a_k in Eq. (8) in terms of b_k and taking the functional derivative in Eq. (23), we obtain

$$\dot{b}_k = -\frac{6\pi i\mu}{\rho c_R \zeta} \int \int_{-\infty}^{\infty} w(k_1, k_2, -k)\delta(k_1+k_2-k)$$
$$\times (b_{k_1} b_{k_2} e^{-i(|k_1|+|k_2|)c_R t} + b_{k_1} b^*_{-k_2} e^{-i(|k_1|-|k_2|)c_R t}$$
$$+ b^*_{-k_1} b_{k_2} e^{i(|k_1|-|k_2|)c_R t}$$
$$+ b^*_{-k_1} b^*_{-k_2} e^{i(|k_1|+|k_2|)c_R t}) e^{i|k|c_R t} \frac{dk_1\, dk_2}{(2\pi)^2}, \qquad (24)$$

where the permutational symmetry of $w(k_1, k_2, k_3)$ was taken into account. Integrating over k_2 gives

$$\dot{b}_k = -\frac{3i\mu}{\rho c_R \zeta} \int_{-\infty}^{\infty} w(k_1, k-k_1, -k)$$
$$\times (b_{k_1} b_{k-k_1} e^{-i(|k_1|+|k-k_1|)c_R t}$$
$$+ b_{k_1} b^*_{k_1-k} e^{-i(|k_1|-|k-k_1|)c_R t}$$
$$+ b^*_{-k_1} b_{k-k_1} e^{i(|k_1|-|k-k_1|)c_R t}$$
$$+ b^*_{-k_1} b^*_{k_1-k} e^{i(|k_1|+|k-k_1|)c_R t}) e^{i|k|c_R t} \frac{dk_1}{2\pi}. \qquad (25)$$

Equation (25) accounts for the interaction of waves that propagate in both the positive and negative x directions, and it can therefore be used to investigate nonlinear effects in compound as well as progressive waves. For example, applications in signal processing involving interactions of counterpropagating waves can be studied with this equation. Equation (25) is thus equivalent to the second-order wave equation.

C. Progressive waves: Evolution in time

Here, approximate equations for progressive plane waves are developed for an initial value problem in which a spatial waveform is prescribed at an instant in time.

To extract from Eq. (25) the evolution equation for progressive waves, we apply the method of averaging.[33] This procedure eliminates the nonresonant nonlinear terms resulting from the interaction of waves that propagate in opposite directions. On the assumption that the nonlinearity is weak, the right-hand side of Eq. (25) becomes small, and the solution for b_k has one part that is a slowly varying function of time, and another small, rapidly oscillating part due to nonresonant interactions. In the first approximation of the method of averaging, rapidly oscillating terms in b_k are eliminated. To separate the resonant and nonresonant terms, we consider each term in Eq. (25) for different values of k_1. The first exponential in Eq. (25) takes the values

$$e^{i(k-|k_1|-|k-k_1|)c_R t} = \begin{cases} e^{2i(k-k_1)c_R t}, & k_1 > k, \\ 1, & 0 < k_1 < k, \\ e^{2ik_1 c_R t}, & k_1 < 0. \end{cases} \qquad (26)$$

Evaluation of the second exponential in Eq. (25) yields

$$e^{i(k-|k_1|+|k-k_1|)c_R t} = \begin{cases} 1, & k_1 > k, \\ e^{2i(k-k_1)c_R t}, & 0 < k_1 < k, \\ e^{2ik c_R t}, & k_1 < 0. \end{cases} \qquad (27)$$

The third exponential becomes

$$e^{i(k+|k_1|-|k-k_1|)c_R t} = \begin{cases} e^{2ik c_R t}, & k_1 > k, \\ e^{2ik_1 c_R t}, & 0 < k_1 < k, \\ 1, & k_1 < 0. \end{cases} \qquad (28)$$

The fourth exponential exhibits an oscillatory behavior for all values of k_1. After averaging Eq. (25) over a large period of time, we are left with

$$\dot{b}_k = -\frac{3i\mu}{\rho c_R \zeta} \left(\int_0^k w(k_1, k-k_1, -k) b_{k_1} b_{k-k_1} \frac{dk_1}{2\pi} \right.$$
$$+ \int_k^{\infty} w(k_1, k-k_1, -k) b_{k_1} b^*_{k_1-k} \frac{dk_1}{2\pi}$$
$$\left. + \int_{-\infty}^0 w(k_1, k-k_1, -k) b^*_{-k_1} b_{k-k_1} \frac{dk_1}{2\pi} \right). \qquad (29)$$

Note that on the right-hand side of Eq. (29), all (now slowly varying) amplitudes b_k and b^*_k have $k>0$, which correspond exclusively to wave propagation in the positive x direction [as seen by substituting Eq. (15) into Eq. (1) or (2)]. In the first approximation of the method of averaging, equations for waves propagating in different directions indeed become separable.

We now introduce the following, more convenient slowly varying spectral amplitudes, which correspond for all k to waves that propagate only in the positive x direction:

$$d_k = \begin{cases} b_k, & k>0, \\ b^*_{-k}, & k<0. \end{cases} \qquad (30)$$

With only forward propagating waves taken into account, Eqs. (1) and (2) can be rewritten as

$$u_x(x,z,t) = \int_{-\infty}^{\infty} i\frac{k}{|k|} d_k(t) h_x(z,k) e^{ik(x-c_R t)} \frac{dk}{2\pi}, \qquad (31)$$

$$u_z(x,z,t) = \int_{-\infty}^{\infty} d_k(t) h_z(z,k) e^{ik(x-c_R t)} \frac{dk}{2\pi}. \qquad (32)$$

When expressed in terms of d_k rather than b_k, the integrals in Eq. (29) may be combined to obtain

$$\dot{d}_k = -\frac{3i\mu k}{\rho c_R \zeta |k|} \int_{-\infty}^{\infty} w(k_1, k-k_1, -k) d_{k_1} d_{k-k_1} \frac{dk_1}{2\pi}. \qquad (33)$$

The solution of Eq. (33), after substitution in Eqs. (31) and (32), describes the evolution in time of a progressive nonlinear Rayleigh wave. Equation (33) is a generalization of Eqs. (33) and (34) in Ref. 1. In contrast with the earlier deriva-

tion, no assumption was made here about periodicity of the signal.

D. Progressive waves: Evolution in space

In practice, we are usually faced with a boundary value problem, i.e., corresponding to radiation from a fixed source. Here, equations describing variations of the spectral amplitudes with distance rather than with time are derived.

We begin by introducing the auxiliary function

$$u(x,t) = \int_{-\infty}^{\infty} d_k(t) e^{ik(x-c_R t)} \frac{dk}{2\pi},$$

(34)

and its Fourier transform in the retarded time frame,

$$\widetilde{u}(x,\omega) = \int_{-\infty}^{\infty} u(x,t) e^{i\omega(t-x/c_R)} dt.$$

(35)

The multiplier $e^{-i\omega x/c_R}$ is included in Eq. (35) to produce a slow dependence of $\widetilde{u}(x,\omega)$ on x. Substituting Eq. (34) into Eq. (35) we obtain

$$\widetilde{u}(x,\omega) = \int_{-\infty}^{\infty} \widetilde{d}_k(\omega-kc_R) e^{-i(\omega-kc_R)x/c_R} \frac{dk}{2\pi},$$

(36)

where

$$\widetilde{d}_k(\omega) = \int_{-\infty}^{\infty} d_k(t) e^{i\omega t} dt.$$

(37)

Equations (36) and (37) relate Fourier transforms in space, $d_k(t)$, and in time, $\widetilde{u}(x,\omega)$, of the displacement function $u(x,t)$. Note that the contribution of $\widetilde{d}_k(\omega)$ to the integral in Eq. (36) is significant only for small values of ω, because d_k varies slowly with time. This fact can be used to express the frequency spectra $\widetilde{u}_x(x,z,\omega)$ and $\widetilde{u}_z(x,z,\omega)$ of the displacement components $u_x(x,z,t)$ and $u_z(x,z,t)$, respectively, in terms of the single function $\widetilde{u}(x,\omega)$.

Calculating the Fourier transform of Eq. (31) with respect to the retarded time frame $t-x/c_R$, we obtain the following expression:

$$\widetilde{u}_x(x,z,\omega) = \int_{-\infty}^{\infty} i \frac{k}{|k|} h_x(z,k) \widetilde{d}_k(\omega-kc_R)$$

$$\times e^{-i(\omega-kc_R)x/c_R} \frac{dk}{2\pi}.$$

(38)

Taking into account that $\widetilde{d}_k(\omega-kc_R)$ contributes significantly to the integral only for $\omega \simeq kc_R$ and that $h_x(z,k)$ varies slowly with k in comparison with $\widetilde{d}_k(\omega-kc_R)$, we can remove h_x from the integral, and making use of Eq. (36) we obtain

$$\widetilde{u}_x(x,z,\omega) = i \frac{\omega}{|\omega|} h_x(z,\omega/c_R) \widetilde{u}(x,\omega).$$

(39)

A similar expression can be obtained for $\widetilde{u}_z(x,z,\omega)$:

$$\widetilde{u}_z(x,z,\omega) = h_z(z,\omega/c_R) \widetilde{u}(x,\omega).$$

(40)

Thus to describe the variations of both $\widetilde{u}_x(x,z,\omega)$ and $\widetilde{u}_z(x,z,\omega)$ in the direction of propagation, it is sufficient to know the governing equation for the single function

$\widetilde{u}(x,\omega)$. The horizontal and vertical displacements can be constructed from $\widetilde{u}(x,\omega)$ as follows:

$$u_x(x,z,t) = \int_{-\infty}^{\infty} i \frac{\omega}{|\omega|} h_x(z,\omega/c_R) \widetilde{u}(x,\omega) e^{i\omega(x/c_R-t)} \frac{d\omega}{2\pi},$$

(41)

$$u_z(x,z,t) = \int_{-\infty}^{\infty} h_z(z,\omega/c_R) \widetilde{u}(x,\omega) e^{i\omega(x/c_R-t)} \frac{d\omega}{2\pi}.$$

(42)

To obtain the equation for the evolution of $\widetilde{u}(x,\omega)$ with distance x, we take the partial derivative of Eq. (36) with respect to x:

$$\frac{\partial \widetilde{u}}{\partial x} = -\frac{i}{c_R} \int_{-\infty}^{\infty} (\omega-kc_R) \widetilde{d}_k(\omega-kc_R)$$

$$\times e^{-i(\omega-kc_R)x/c_R} \frac{dk}{2\pi}.$$

(43)

Noting that $-i\omega \widetilde{d}_k(\omega)$ is the Fourier transform of the time derivative of $d_k(t)$, we can replace Eq. (43) by

$$\frac{\partial \widetilde{u}}{\partial x} = \frac{1}{c_R} \int_{-\infty}^{\infty} \widetilde{\dot{d}}_k(\omega-kc_R) e^{-i(\omega-kc_R)x/c_R} \frac{dk}{2\pi},$$

(44)

where $\widetilde{\dot{d}}_k(\omega)$ is the Fourier transform of $\dot{d}_k(t)$. Taking the Fourier transform of Eq. (33), we obtain

$$\widetilde{\dot{d}}_k(\omega-kc_R) = -\frac{6i\pi\mu k}{\rho c_R \zeta |k|} \int \int_{-\infty}^{\infty} \frac{dk_1\, dk_2}{(2\pi)^2} w(k_1,k_2,-k)$$

$$\times \delta(k-k_1-k_2) \int_{-\infty}^{\infty} dt\, d_{k_1}(t) d_{k_2}(t)$$

$$\times e^{it(\omega-kc_R)}.$$

(45)

After Fourier integrals are substituted in place of $d_{k_1}(t)$ and $d_{k_2}(t)$, the integration over time can be performed:

$$\widetilde{\dot{d}}_k(\omega-kc_R) = -\frac{12i\pi^2\mu k}{\rho c_R \zeta |k|} \int \int_{-\infty}^{\infty} \frac{dk_1\, dk_2}{(2\pi)^2} w(k_1,k_2,-k)$$

$$\times \delta(k-k_1-k_2) \int \int_{-\infty}^{\infty} \frac{d\omega_1\, d\omega_2}{(2\pi)^2}$$

$$\times \delta(\omega-\omega_1-\omega_2-kc_R) \widetilde{d}_{k_1}(\omega_1) \widetilde{d}_{k_2}(\omega_2). \quad (46)$$

Letting $\omega_1' = \omega_1+k_1 c_R$ and $\omega_2' = \omega_2+k_2 c_R$, and taking into account that $k=k_1+k_2$, we can write Eq. (46) in the form

$$\widetilde{\dot{d}}_k(\omega-kc_R) = -\frac{12i\pi^2\mu k}{\rho c_R \zeta |k|} \int \int_{-\infty}^{\infty} \frac{dk_1\, dk_2}{(2\pi)^2} w(k_1,k_2,-k)$$

$$\times \delta(k-k_1-k_2) \int \int_{-\infty}^{\infty} \frac{d\omega_1\, d\omega_2}{(2\pi)^2}$$

$$\times \delta(\omega-\omega_1-\omega_2) \widetilde{d}_{k_1}(\omega_1-k_1 c_R)$$

$$\times \widetilde{d}_{k_2}(\omega_2-k_2 c_R).$$

(47)

By substituting Eq. (47) into Eq. (44) and integrating over k, we obtain

1407   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Knight *et al.*: Nonlinear Rayleigh wave theory   1407

$$\frac{\partial \widetilde{u}}{\partial x} = -\frac{6i\pi\mu}{\rho c_R^2 \zeta} \int \int \int \int_{-\infty}^{\infty} \frac{k_1+k_2}{|k_1+k_2|} \delta(\omega-\omega_1-\omega_2)$$

$$\times w(k_1,k_2,-k_1-k_2)\widetilde{d}_{k_1}(\omega_1-k_1c_R)$$

$$\times \widetilde{d}_{k_2}(\omega_2-k_2c_R)$$

$$\times e^{i(k_1+k_2-\omega/c_R)x} \frac{d\omega_1 \, d\omega_2}{(2\pi)^2} \frac{dk_1 \, dk_2}{(2\pi)^2}. \tag{48}$$

Since $\widetilde{d}_k(\omega-kc_R)$ contributes significantly only for $\omega \simeq kc_R$, the main contribution to the integrals over k_1 and k_2 is obtained from wave numbers that are close to ω_1/c_R and ω_2/c_R, respectively. Therefore, Eq. (48) can be approximated by

$$\frac{\partial \widetilde{u}}{\partial x} = -\frac{6i\pi\mu\omega}{\rho c_R^2 \zeta |\omega|} \int \int_{-\infty}^{\infty} \frac{d\omega_1 \, d\omega_2}{(2\pi)^2} \delta(\omega-\omega_1-\omega_2)$$

$$\times w(\omega_1/c_R,\omega_2/c_R,-\omega/c_R) \int \int_{-\infty}^{\infty} \frac{dk_1 \, dk_2}{(2\pi)^2}$$

$$\times \widetilde{d}_{k_1}(\omega_1-k_1c_R)\widetilde{d}_{k_2}(\omega_2-k_2c_R)e^{i(k_1+k_2-\omega/c_R)x}, \tag{49}$$

where the fact that $w(k_1,k_2,-k_1-k_2)$ varies slowly with k_1 and k_2 in comparison with $\widetilde{d}_k(\omega-kc_R)$ was used. Recognizing that the integrals over k_1 and k_2 coincide with Eq. (36), we can reduce Eq. (49) to

$$\frac{\partial \widetilde{u}}{\partial x} = -\frac{6i\pi\mu\omega}{\rho c_R^2 \zeta |\omega|} \int \int_{-\infty}^{\infty} \delta(\omega-\omega_1-\omega_2)$$

$$\times w(\omega_1/c_R,\omega_2/c_R,-\omega/c_R)\widetilde{u}(x,\omega_1)$$

$$\times \widetilde{u}(x,\omega_2) \frac{d\omega_1 \, d\omega_2}{(2\pi)^2}. \tag{50}$$

Integration with respect to ω_2 yields

$$\frac{\partial \widetilde{u}}{\partial x} = -\frac{3i\mu\omega}{\rho c_R^4 \zeta |\omega|} \int_{-\infty}^{\infty} w(\omega_1,\omega-\omega_1,-\omega)$$

$$\times \widetilde{u}(x,\omega_1)\widetilde{u}(x,\omega-\omega_1) \frac{d\omega_1}{2\pi}, \tag{51}$$

where the relation $w(\omega_1/c_R,\omega_2/c_R,\omega_3/c_R)=w(\omega_1,\omega_2,\omega_3)/c_R^2$, which follows from Eq. (9), was used.

We now let

$$R(\omega_1,\omega_2) = -\frac{3}{|\omega_1\omega_2\omega|} w(\omega_1,\omega_2,-\omega), \tag{52}$$

where $\omega=\omega_1+\omega_2$. Substitution of Eq. (9) yields

$$R(\omega_1,\omega_2) = \frac{\alpha'}{|\omega_1+\omega_2|\xi_l+|\omega_1|\xi_t+|\omega_2|\xi_t} + \frac{\alpha'}{|\omega_1+\omega_2|\xi_t+|\omega_1|\xi_t+|\omega_2|\xi_l} + \frac{\alpha'}{|\omega_1+\omega_2|\xi_t+|\omega_1|\xi_l+|\omega_2|\xi_t}$$

$$+ \frac{\beta'}{|\omega_1+\omega_2|\xi_t+|\omega_1|\xi_l+|\omega_2|\xi_l} + \frac{\beta'}{|\omega_1+\omega_2|\xi_l+|\omega_1|\xi_t+|\omega_2|\xi_l} + \frac{\beta'}{|\omega_1+\omega_2|\xi_l+|\omega_1|\xi_l+|\omega_2|\xi_t}$$

$$+ \frac{3\gamma'}{|\omega_1+\omega_2|\xi_l+|\omega_1|\xi_l+|\omega_2|\xi_l}. \tag{53}$$

When ω_1 and ω_2 are integer multiples of some frequency ω_0 ($\omega_1=m\omega_0$, $\omega_2=l\omega_0$), $R(\omega_1,\omega_2)$ reduces to R_{ml}/ω_0, where R_{ml} is the matrix that has been used in previous work.[29] Equation (51) thus assumes the form

$$\frac{\partial \widetilde{u}}{\partial x} = \frac{i\mu\omega}{\rho c_R^4 \zeta} \int_{-\infty}^{\infty} |\omega'(\omega-\omega')|R(\omega',\omega-\omega')$$

$$\times \widetilde{u}(x,\omega')\widetilde{u}(x,\omega-\omega') \frac{d\omega'}{2\pi}, \tag{54}$$

which is the model equation for the function $\widetilde{u}(x,\omega)$. As shown in Appendix A, Eq. (54) is equivalent to the spectral evolution equation derived by Parker.[2]

The set of coupled equations derived previously in Ref. 1 is for spectral amplitudes of the velocity field. To rewrite the governing equations in terms of the velocity spectrum we use the relation $\widetilde{v}_j=-i\omega\widetilde{u}_j$ (where $j=x,z$), and the spectral velocity function $\widetilde{v}(x,\omega)=-i\omega\widetilde{u}(x,\omega)$ is introduced. The velocity components may then be expressed as

$$v_x(x,z,t) = \int_{-\infty}^{\infty} i\frac{\omega}{|\omega|} h_x(z,\omega/c_R)\widetilde{v}(x,\omega)e^{i\omega(x/c_R-t)} \frac{d\omega}{2\pi}, \tag{55}$$

$$v_z(x,z,t) = \int_{-\infty}^{\infty} h_z(z,\omega/c_R)\widetilde{v}(x,\omega)e^{i\omega(x/c_R-t)} \frac{d\omega}{2\pi}. \tag{56}$$

With Eq. (54) multiplied by $-i\omega$, the equation for $\widetilde{v}(x,\omega)$ assumes the form

$$\frac{\partial \widetilde{v}}{\partial x} = -\frac{\mu\omega^2}{\rho c_R^4 \zeta} \int_{-\infty}^{\infty} \frac{\omega'(\omega-\omega')}{|\omega'(\omega-\omega')|}$$

$$\times R(\omega',\omega-\omega')\widetilde{v}(x,\omega')\widetilde{v}(x,\omega-\omega') \frac{d\omega'}{2\pi}. \tag{57}$$

In contrast to Eq. (38) in Ref. 1, Eq. (57) above is not restricted to periodic waveforms. The discrete form of Eq. (57) is given by Eq. (B8) with $s=0$.

II. NONPLANAR WAVES

A spectral evolution equation is derived below in Sec. II C for nonlinear propagation of nonplanar progressive waves for cases in which noncumulative nonlinear effects are not of great importance. Noncumulative nonlinear effects are associated with noncollinear interactions of plane-wave components in the angular spectrum of the field. An approximation is introduced whereby the nonlinear term is correct to second order only for interacting components of the angular spectrum that propagate in the same direction. The resulting equation provides an accurate model when the dominant nonlinear effects are due to components of the angular spectrum that interact at small angles. It can therefore be used to model progressive waves such as cylindrical waves for which $kr \gg 1$ (where r is the radius of curvature of the wavefront), and directive beams for which $ka \gg 1$ (where a characterizes the transverse dimension of the source). It cannot be used to model fields in which the main nonlinear effects of interest result from components of the angular spectrum that interact at large angles. In this regard, the limitations on the equation obtained in Sec. II C are similar to those that apply to the Westervelt equation[34] used in nonlinear acoustics.

A. Basic equations

The coordinate system is the same as in Sec. I, but the y dependence of the field variables must now be taken into account. The wave vector $\mathbf{k}$ of each harmonic component is parallel to the (x,y) plane and has two components, k_x and k_y. The displacement vector has three components, u_x, u_y, and u_z. It is convenient to define a two-dimensional displacement vector $\mathbf{u}$ in the (x,y) plane, with components (u_x,u_y), and a vector $\mathbf{r}$ parallel to the (x,y) plane. Since the derivation for plane waves was discussed in detail in Sec. I, only the differences shall be emphasized here.

We start with a Fourier expansion of the displacement field, but now two transforms are taken. The expressions can be easily obtained from Eqs. (1) and (2) by replacing x by vector $\mathbf{r}$, and k by $\mathbf{k}$. The two-dimensional Fourier decomposition of the displacement field is given by

$$\mathbf{u}(\mathbf{r},z,t) = \int i \frac{\mathbf{k}}{|\mathbf{k}|} a_{\mathbf{k}}(t) h_{\perp}(z,|\mathbf{k}|) e^{i\mathbf{k}\cdot\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^2}, \quad (58)$$

$$u_z(\mathbf{r},z,t) = \int a_{\mathbf{k}}(t) h_z(z,|\mathbf{k}|) e^{i\mathbf{k}\cdot\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^2}, \quad (59)$$

where we employ the notation $d\mathbf{k} = dk_x \, dk_y$. The depth-dependent functions $h_{\perp}(z,|\mathbf{k}|)$ and $h_z(z,|\mathbf{k}|)$ are the same as for plane waves and given by

$$h_{\perp}(z,|\mathbf{k}|) = \xi_t e^{|\mathbf{k}|\xi_t z} + \eta e^{|\mathbf{k}|\xi_l z}, \quad (60)$$

$$h_z(z,|\mathbf{k}|) = e^{|\mathbf{k}|\xi_t z} + \eta \xi_l e^{|\mathbf{k}|\xi_l z}. \quad (61)$$

Expressions analogous to Eqs. (5), (7), and (8) for the kinetic, quadratic and cubic elastic energies (here, all have dimensions of energy) take the form

$$T = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{0} \frac{\rho}{2} (\dot{u}_x^2 + \dot{u}_y^2 + \dot{u}_z^2) dx \, dy \, dz$$

$$= \frac{\rho\zeta}{4} \int \frac{\dot{a}_{\mathbf{k}}\dot{a}_{-\mathbf{k}}}{|\mathbf{k}|} \frac{d\mathbf{k}}{(2\pi)^2}, \quad (62)$$

$$V = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{0} \mathscr{E}_2 \, dx \, dy \, dz$$

$$= \frac{\rho c_R^2 \zeta}{4} \int a_{\mathbf{k}} a_{-\mathbf{k}} |\mathbf{k}| \frac{d\mathbf{k}}{(2\pi)^2}, \quad (63)$$

$$W = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{0} \mathscr{E}_3 \, dx \, dy \, dz$$

$$= (2\pi)^2 \mu \int \int \int w(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) a_{\mathbf{k}_1} a_{\mathbf{k}_2} a_{\mathbf{k}_3}$$

$$\times \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \frac{d\mathbf{k}_1 \, d\mathbf{k}_2 \, d\mathbf{k}_3}{(2\pi)^6}. \quad (64)$$

The expression for $w(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ is different from Eq. (9) for $w(k_1,k_2,k_3)$. The latter corresponds to the case where all wave vectors participating in the interaction are parallel. The general equation for $w(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ could be obtained explicitly by direct calculation of the integrals involved in the interaction energy. However, in many practical cases of interest the dominant nonlinear effects are produced by components in the angular spectrum that interact at small angles, for which use of the expression $w(k_1,k_2,k_3)$ is a reasonable approximation. In such cases, there is no need to evaluate the general expression $w(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$.

The system Hamiltonian is given again by Eq. (11).

B. Hamiltonian formalism

Spectral densities $a_{\mathbf{k}}$ are taken for generalized coordinates, and the corresponding canonical momenta are given by

$$p_{\mathbf{k}} = \frac{\delta T}{\delta \dot{a}_{\mathbf{k}}} = \frac{\rho\zeta}{8\pi^2 |\mathbf{k}|} \dot{a}_{-\mathbf{k}}. \quad (65)$$

Expressing $\dot{a}_{\mathbf{k}}$ in Eq. (62) in terms of $p_{-\mathbf{k}}$, we obtain

$$T = \frac{(2\pi)^4}{\rho\zeta} \int p_{\mathbf{k}} p_{-\mathbf{k}} |\mathbf{k}| \frac{d\mathbf{k}}{(2\pi)^2}. \quad (66)$$

The Hamiltonian canonical equations are the same as for plane waves and given by Eqs. (14), but with k replaced by $\mathbf{k}$. New variables $b_{\mathbf{k}}$ and $b_{\mathbf{k}}^*$ are defined by

$$a_{\mathbf{k}}(t) = b_{\mathbf{k}}(t) e^{-i|\mathbf{k}|c_R t} + b_{-\mathbf{k}}^*(t) e^{i|\mathbf{k}|c_R t}, \quad (67)$$

$$\dot{a}_{\mathbf{k}}(t) = -i|\mathbf{k}|c_R b_{\mathbf{k}}(t) e^{-i|\mathbf{k}|c_R t} + i|\mathbf{k}|c_R b_{-\mathbf{k}}^*(t) e^{i|\mathbf{k}|c_R t}. \quad (68)$$

In terms of the new amplitudes $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}^*$, the Hamiltonian can be written as

$$H = \rho c_R^2 \zeta \int b_{\mathbf{k}} b_{\mathbf{k}}^* |k| \frac{d\mathbf{k}}{(2\pi)^2} + W. \quad (69)$$

Equations (67) and (68) are inverted, and Eq. (65) is used to express the amplitudes $b_{\mathbf{k}}$ in terms of the canonical variables:

$$b_{\mathbf{k}} = \left(\frac{a_{\mathbf{k}}}{2} + i \frac{4\pi^2}{\rho c_R \zeta} p_{-\mathbf{k}} \right) e^{i|\mathbf{k}|c_R t}. \tag{70}$$

The derivative of Eq. (70) with respect to time, combined with Eqs. (14) and (69), yields

$$\dot{b}_{\mathbf{k}} = -\frac{4i\pi^2}{\rho c_R \zeta} \frac{\delta W}{\delta b_{\mathbf{k}}^*}. \tag{71}$$

Equation (71) is a generalization of Eq. (23).

After substituting Eq. (64) into Eq. (71), where $a_{\mathbf{k}}$ is replaced by $b_{\mathbf{k}}$ according to Eq. (67), and taking the functional derivative, we obtain

$$\dot{b}_{\mathbf{k}} = -\frac{12i\pi^2\mu}{\rho c_R \zeta} \int \int w(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}) \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k})$$

$$\times (b_{\mathbf{k}_1} b_{\mathbf{k}_2} e^{-i(|\mathbf{k}_1|+|\mathbf{k}_2|)c_R t} + b_{\mathbf{k}_1} b_{-\mathbf{k}_2}^* e^{-i(|\mathbf{k}_1|-|\mathbf{k}_2|)c_R t}$$

$$+ b_{-\mathbf{k}_1}^* b_{\mathbf{k}_2} e^{i(|\mathbf{k}_1|-|\mathbf{k}_2|)c_R t}$$

$$+ b_{-\mathbf{k}_1}^* b_{-\mathbf{k}_2}^* e^{i(|\mathbf{k}_1|+|\mathbf{k}_2|)c_R t}) e^{i|\mathbf{k}|c_R t} \frac{d\mathbf{k}_1 \, d\mathbf{k}_2}{(2\pi)^4}. \tag{72}$$

Equation (72) is correct to quadratic order (again, subject to the assumption that surface waves do not interact with bulk waves). It is very similar to Eq. (24), and it can be used to investigate nonlinear distortion in wave fields with arbitrary spatial dependencies (e.g., compound waves), but it requires calculation of the function $w(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$.

C. Progressive waves: Evolution in space

We now assume that the nonlinearity is weak. As was discussed in Sec. I C, the solution for $b_{\mathbf{k}}$ has a dominant, slowly varying part and a small, rapidly oscillating part. It can be seen from Eq. (72) that only harmonic components with wave vectors $\mathbf{k}$ that are nearly collinear contribute to the slowly varying part of the amplitudes $b_{\mathbf{k}}$. This fact is due to the nonresonant character of noncollinear interactions. An approximation may therefore be made at this point, namely, replacement of the coupling coefficient $w(\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k})$ for the interaction of waves with wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ with the coupling coefficient $w[|\mathbf{k}_1|, |\mathbf{k}_2|, -(|\mathbf{k}_1| \pm |\mathbf{k}_2|)]$, which corresponds to collinear interactions. The ''+'' sign is taken with the terms $b_{\mathbf{k}_1} b_{\mathbf{k}_2}$ and $b_{-\mathbf{k}_1}^* b_{-\mathbf{k}_2}^*$, and the ''−'' sign corresponds to $b_{\mathbf{k}_1} b_{-\mathbf{k}_2}^*$ and $b_{-\mathbf{k}_1}^* b_{\mathbf{k}_2}$, so that the leading interactions of waves that propagate in the same direction are included with the correct coupling coefficient. As a result, all noncollinear interactions still contribute to the right-hand side of Eq. (72), but without the correct coupling coefficient. The obtained approximate results will be identical to the solutions of Eq. (72) to first order in the method of averaging. After the substitutions described above are made, Eq. (72) takes the form

$$\dot{b}_{\mathbf{k}} = -\frac{12i\pi^2\mu}{\rho c_R \zeta} \int \int \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k})[w(|\mathbf{k}_1|, |\mathbf{k}_2|,$$

$$-|\mathbf{k}_1| - |\mathbf{k}_2|)(b_{\mathbf{k}_1} b_{\mathbf{k}_2} e^{-i(|\mathbf{k}_1|+|\mathbf{k}_2|)c_R t}$$

$$+ b_{-\mathbf{k}_1}^* b_{-\mathbf{k}_2}^* e^{i(|\mathbf{k}_1|+|\mathbf{k}_2|)c_R t}) + w(|\mathbf{k}_1|, |\mathbf{k}_2|,$$

$$-|\mathbf{k}_1| + |\mathbf{k}_2|)(b_{\mathbf{k}_1} b_{-\mathbf{k}_2}^* e^{-i(|\mathbf{k}_1|-|\mathbf{k}_2|)c_R t}$$

$$+ b_{-\mathbf{k}_1}^* b_{\mathbf{k}_2} e^{i(|\mathbf{k}_1|-|\mathbf{k}_2|)c_R t})] e^{i|\mathbf{k}|c_R t} \frac{d\mathbf{k}_1 \, d\mathbf{k}_2}{(2\pi)^4}. \tag{73}$$

An important comment at this point is that in contrast with the derivation in Sec. I for the case of plane waves, no time averaging was performed. Equation (73) still contains contributions due to the nonresonant noncollinear interactions, although they are included with incorrect coupling coefficients.

We now reformulate the initial value problem as a boundary value problem by relating the Fourier transforms in space and time. To reconstruct the displacement field, substitute Eq. (67) into Eqs. (58) and (59) to obtain

$$\mathbf{u}(\mathbf{r},z,t) = \int i \frac{\mathbf{k}}{|\mathbf{k}|} [b_{\mathbf{k}}(t)e^{-i|\mathbf{k}|c_R t} + b_{-\mathbf{k}}^*(t)e^{i|\mathbf{k}|c_R t}]$$

$$\times h_\perp(z,|\mathbf{k}|)e^{i\mathbf{k}\cdot\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^2}, \tag{74}$$

$$u_z(\mathbf{r},z,t) = \int [b_{\mathbf{k}}(t)e^{-i|\mathbf{k}|c_R t} + b_{-\mathbf{k}}^*(t)e^{i|\mathbf{k}|c_R t}]$$

$$\times h_z(z,|\mathbf{k}|)e^{i\mathbf{k}\cdot\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^2}. \tag{75}$$

By analogy with Eq. (34), a function $U(\mathbf{r},t)$ is defined as

$$U(\mathbf{r},t) = \int [b_{\mathbf{k}}(t)e^{-i|\mathbf{k}|c_R t} + b_{-\mathbf{k}}^*(t)e^{i|\mathbf{k}|c_R t}]e^{i\mathbf{k}\cdot\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^2}. \tag{76}$$

The temporal Fourier transform of $U(\mathbf{r},t)$, defined as

$$\widetilde{U}(\mathbf{r},\omega) = \int U(\mathbf{r},t)e^{i\omega t} \, dt, \tag{77}$$

becomes

$$\widetilde{U}(\mathbf{r},\omega) = \int [\widetilde{b}_{\mathbf{k}}(\omega - |\mathbf{k}|c_R)$$

$$+ \widetilde{b}_{-\mathbf{k}}^*(\omega + |\mathbf{k}|c_R)]e^{i\mathbf{k}\cdot\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^2}, \tag{78}$$

where $\widetilde{b}_{\mathbf{k}}(\omega)$ is the Fourier transform of $b_{\mathbf{k}}(t)$. The frequency spectra $\widetilde{\mathbf{u}}(\mathbf{r},z,\omega)$ and $\widetilde{u}_z(\mathbf{r},z,\omega)$ of the displacement components $\mathbf{u}(\mathbf{r},z,t)$ and $u_z(\mathbf{r},z,t)$ can be expressed in terms of the function $\widetilde{U}(\mathbf{r},\omega)$ in a form similar to Eqs. (39) and (40):

$$\widetilde{\mathbf{u}}(\mathbf{r},z,\omega) = \frac{c_R}{|\omega|} h_\perp(z,\omega/c_R) \mathbf{\nabla}_\perp \widetilde{U}(\mathbf{r},\omega), \qquad (79)$$

$$\widetilde{u}_z(\mathbf{r},z,\omega) = h_z(z,\omega/c_R) \widetilde{U}(\mathbf{r},\omega), \qquad (80)$$

where $\mathbf{\nabla}_\perp$ is the gradient operator in the (x,y) plane.

The governing equation for the function $\widetilde{U}(\mathbf{r},\omega)$ can be obtained by applying the operator $(\nabla_\perp^2 + \omega^2/c_R^2)$ to the functional relation between Fourier transforms in time and in space given by Eq. (78):

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U}(\mathbf{r},\omega) = \frac{1}{c_R^2} \int e^{i\mathbf{k}\cdot\mathbf{r}} (\omega - |\mathbf{k}|c_R)(\omega + |\mathbf{k}|c_R)$$

$$\times [\widetilde{b}_\mathbf{k}(\omega - |\mathbf{k}|c_R)$$

$$+ \widetilde{b}^*_{-\mathbf{k}}(\omega + |\mathbf{k}|c_R)] \frac{d\mathbf{k}}{(2\pi)^2}. \qquad (81)$$

Replacing $-i\omega \widetilde{b}_\mathbf{k}(\omega)$ by $\widetilde{\dot{b}}_\mathbf{k}(\omega)$, where $\widetilde{\dot{b}}_\mathbf{k}(\omega)$ is the Fourier transform of $\dot{b}_\mathbf{k}(t)$, we obtain

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U} = \frac{i}{c_R^2} \int e^{i\mathbf{k}\cdot\mathbf{r}} [(\omega + |\mathbf{k}|c_R)\widetilde{b}_\mathbf{k}(\omega - |\mathbf{k}|c_R)$$

$$+ (\omega - |\mathbf{k}|c_R)\widetilde{b}^*_{-\mathbf{k}}(\omega + |\mathbf{k}|c_R)] \frac{d\mathbf{k}}{(2\pi)^2}. \qquad (82)$$

Recall that $\widetilde{b}_\mathbf{k}(\omega)$ makes a significant contribution to the integral only for small values of ω because the dominant component of $b_\mathbf{k}(t)$ varies slowly with time. Equation (82) thus reduces to

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U} = \frac{2i\omega}{c_R^2} \int e^{i\mathbf{k}\cdot\mathbf{r}} [\widetilde{b}_\mathbf{k}(\omega - |\mathbf{k}|c_R)$$

$$+ \widetilde{b}^*_{-\mathbf{k}}(\omega + |\mathbf{k}|c_R)] \frac{d\mathbf{k}}{(2\pi)^2}. \qquad (83)$$

Now consider the positive frequencies $\omega > 0$, for which only the term with $\widetilde{b}_\mathbf{k}(\omega - |\mathbf{k}|c_R)$ is important. Using Eq. (73), we obtain

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U} = \frac{48\pi^3\omega\mu}{\rho c_R^3 \zeta} \int \int \int \int \int e^{i\mathbf{k}\cdot\mathbf{r}} \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}) \delta(\omega_1 + \omega_2 - \omega)\{w(|\mathbf{k}_1|,|\mathbf{k}_2|,-|\mathbf{k}_1| - |\mathbf{k}_2|)[\widetilde{b}_{\mathbf{k}_1}(\omega_1 - |\mathbf{k}_1|c_R)$$

$$\times \widetilde{b}_{\mathbf{k}_2}(\omega_2 - |\mathbf{k}_2|c_R) + \widetilde{b}^*_{-\mathbf{k}_1}(\omega_1 + |\mathbf{k}_1|c_R)\widetilde{b}^*_{-\mathbf{k}_2}(\omega_2 + |\mathbf{k}_2|c_R)] + w(|\mathbf{k}_1|,|\mathbf{k}_2|,-|\mathbf{k}_1| + |\mathbf{k}_2|)[\widetilde{b}_{\mathbf{k}_1}(\omega_1 - |\mathbf{k}_1|c_R)$$

$$\times \widetilde{b}^*_{-\mathbf{k}_2}(\omega_2 + |\mathbf{k}_2|c_R) + \widetilde{b}^*_{-\mathbf{k}_1}(\omega_1 + |\mathbf{k}_1|c_R)\widetilde{b}_{\mathbf{k}_2}(\omega_2 - |\mathbf{k}_2|c_R)]\} \frac{d\mathbf{k}}{(2\pi)^2} \frac{d\mathbf{k}_1\, d\mathbf{k}_2}{(2\pi)^4} \frac{d\omega_1\, d\omega_2}{(2\pi)^2}, \qquad (84)$$

where the following relations were used:

$$\widetilde{b}_\mathbf{k}(\omega - |\mathbf{k}|c_R) = \int \dot{b}_\mathbf{k}(t) e^{i(\omega - |\mathbf{k}|c_R)t}\, dt,$$

$$b_\mathbf{k}(t) e^{-i|\mathbf{k}|c_R t} = \int \widetilde{b}_\mathbf{k}(\omega - |\mathbf{k}|c_R) e^{-i\omega t} \frac{d\omega}{2\pi},$$

$$b^*_{-\mathbf{k}}(t) e^{i|\mathbf{k}|c_R t} = \int \widetilde{b}^*_{-\mathbf{k}}(\omega + |\mathbf{k}|c_R) e^{-i\omega t} \frac{d\omega}{2\pi}.$$

The integration over $\mathbf{k}$ in Eq. (84) can now be performed. Taking into account that $\widetilde{b}_\mathbf{k}(\omega)$ is significant only for small ω, and that $w(|\mathbf{k}_1|,|\mathbf{k}_2|,|\mathbf{k}_3|)$ varies slowly as a function of its arguments, Eq. (84) can be written in the form

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U} = \frac{12\pi\omega\mu}{\rho c_R^3 \zeta} \int \int \frac{d\omega_1\, d\omega_2}{(2\pi)^2} \delta(\omega_1 + \omega_2 - \omega) w[\omega_1/c_R, \omega_2/c_R, -(\omega_1 + \omega_2)/c_R]$$

$$\times \int \int \frac{d\mathbf{k}_1\, d\mathbf{k}_2}{(2\pi)^4} e^{i(\mathbf{k}_1 + \mathbf{k}_2)\cdot\mathbf{r}} [\widetilde{b}_{\mathbf{k}_1}(\omega_1 - |\mathbf{k}_1|c_R) + \widetilde{b}^*_{-\mathbf{k}_1}(\omega_1 + |\mathbf{k}_1|c_R)]$$

$$\times [\widetilde{b}_{\mathbf{k}_2}(\omega_2 - |\mathbf{k}_2|c_R) + \widetilde{b}^*_{-\mathbf{k}_2}(\omega_2 + |\mathbf{k}_2|c_R)]. \qquad (85)$$

Substitution of Eq. (78) into Eq. (85) yields

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U} = \frac{12\pi\omega\mu}{\rho c_R^3 \zeta}\int\int w[\omega_1/c_R, \omega_2/c_R,$$

$$-(\omega_1+\omega_2)/c_R]\delta(\omega_1+\omega_2-\omega)$$

$$\times \widetilde{U}(\mathbf{r},\omega_1)\widetilde{U}(\mathbf{r},\omega_2)\frac{d\omega_1\,d\omega_2}{(2\pi)^2}. \quad (86)$$

Repeating the same procedure for $\omega < 0$, we obtain

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U} = -\frac{12\pi\omega\mu}{\rho c_R^3 \zeta}\int\int w[\omega_1/c_R, \omega_2/c_R,$$

$$-(\omega_1+\omega_2)/c_R]\delta(\omega_1+\omega_2-\omega)$$

$$\times \widetilde{U}(\mathbf{r},\omega_1)\widetilde{U}(\mathbf{r},\omega_2)\frac{d\omega_1\,d\omega_2}{(2\pi)^2}. \quad (87)$$

Equations (86) and (87) can be combined into a single equation that is valid for all ω:

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{U} = \frac{12\pi|\omega|\mu}{\rho c_R^5 \zeta}\int\int w(\omega_1,\omega_2,-\omega_1-\omega_2)$$

$$\times \delta(\omega_1+\omega_2-\omega)$$

$$\times \widetilde{U}(\mathbf{r},\omega_1)\widetilde{U}(\mathbf{r},\omega_2)\frac{d\omega_1\,d\omega_2}{(2\pi)^2}, \quad (88)$$

which describes the spatial evolution of the spectral density $\widetilde{U}(\mathbf{r},\omega)$.

Equation (88) is the nonlinear Helmholtz equation for surface wave fields with arbitrary time and space dependencies. It is correct to first order in the method of averaging. Equation (88) is analogous to the spectral form of the Westervelt equation[34] used in nonlinear acoustics, insofar as the left-hand side accounts for all diffraction effects in the linear approximation, and the nonlinear term on the right-hand side is correct to second order only for the collinear interaction of plane waves. In other words, cumulative nonlinear effects which result from collinear interactions are included correctly to $O(\epsilon^2)$, whereas noncumulative effects which result from noncollinear interactions are taken into account with incorrect coupling coefficients. As noted in the introduction to Sec. II, Eq. (88) provides an accurate model of fields in which the dominant nonlinear effects result from components of the angular spectrum that interact at small angles, e.g., in cylindrical waves for which $kr \gg 1$, and in directive beams for which $ka \gg 1$.

D. Spatial evolution of the velocity

The particle velocity spectrum is related to the particle displacement spectrum according to $\widetilde{v}_j(\mathbf{r},z,\omega) = -i\omega\widetilde{u}_j(\mathbf{r},z,\omega)$, with $j = x,y,z$. We also introduce the function $\widetilde{V}(\mathbf{r},\omega) = -i\omega\widetilde{U}(\mathbf{r},\omega)$ to obtain

$$\mathbf{v}(\mathbf{r},z,t) = i\int_{-\infty}^{\infty} \frac{c_R}{\omega} h_\perp(z,\omega/c_R)\nabla_\perp\widetilde{V}(\mathbf{r},\omega)e^{-i\omega t}\frac{d\omega}{2\pi}, \quad (89)$$

$$v_z(\mathbf{r},z,t) = \int_{-\infty}^{\infty} h_z(z,\omega/c_R)\widetilde{V}(\mathbf{r},\omega)e^{-i\omega t}\frac{d\omega}{2\pi}. \quad (90)$$

The equation for $\widetilde{V}(\mathbf{r},\omega)$ is obtained by multiplying Eq. (88) by $-i\omega$:

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{V} = \frac{12i\pi|\omega|\omega\mu}{\rho c_R^5 \zeta}\int\int\frac{1}{\omega_1\omega_2}w(\omega_1,\omega_2,$$

$$-\omega_1-\omega_2)\delta(\omega_1+\omega_2-\omega)$$

$$\times \widetilde{V}(\mathbf{r},\omega_1)\widetilde{V}(\mathbf{r},\omega_2)\frac{d\omega_1\,d\omega_2}{(2\pi)^2}. \quad (91)$$

Performing the integration over ω_2 yields

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{V} = \frac{6i|\omega|\omega\mu}{\rho c_R^5 \zeta}\int\frac{w(\omega',\omega-\omega',-\omega)}{\omega'(\omega-\omega')}$$

$$\times \widetilde{V}(\mathbf{r},\omega')\widetilde{V}(\mathbf{r},\omega-\omega')\frac{d\omega'}{2\pi}. \quad (92)$$

Finally, use of Eq. (52) permits Eq. (92) to be written in the form

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c_R^2}\right)\widetilde{V} = -\frac{2i\omega^3\mu}{\rho c_R^5 \zeta}\int_{-\infty}^{\infty}\frac{\omega'(\omega-\omega')}{|\omega'(\omega-\omega')|}R(\omega',\omega-\omega')$$

$$\times \widetilde{V}(\mathbf{r},\omega')\widetilde{V}(\mathbf{r},\omega-\omega')\frac{d\omega'}{2\pi}. \quad (93)$$

The remarks in the paragraph following Eq. (88) apply also to Eq. (93). In Appendix B, Eq. (93) is reduced to discrete form for a periodic signal.

E. Cylindrical waves

Here we simplify Eq. (93) for application to progressive cylindrical waves (and progressive plane waves as a degenerate case) by introducing appropriate slow scales.

Let all field variables depend on the radial coordinate r, in which case the displacement vector in the (x,y) plane has only the u_r component. We seek a consistent equation on the slow scale, i.e., in the retarded time frame $t' = t - (r-r_0)/c_R$, where r_0 is the radius of the source. The Fourier transform of the velocity function in the retarded time frame is related to $\widetilde{V}(r,\omega)$ by

$$\widetilde{v}(r,\omega) = e^{-i\omega(r-r_0)/c_R}\widetilde{V}(r,\omega). \quad (94)$$

The model equation for nonplanar waves, Eq. (93), was obtained in the quadratic approximation, i.e., to $O(\epsilon^2)$, where $\epsilon = u_0/L$. The radius of the source that generates the cylindrical wave is assumed to exceed the maximum wavelength of the radiated signal (in order that $kr \gg 1$ is satisfied throughout the field, to ensure that the dominant nonlinear effects are due to interactions of waves that are nearly collinear). Nonlinear effects cause a gradual distortion of waveforms in the retarded time frame over a slow scale measured by $r_1 = \epsilon r$. Rewriting Eq. (93) for $\widetilde{v}(r,\omega)$ and making the coordinate transformation from r to r_1, we obtain

$$\left(\epsilon^2 \frac{\partial^2}{\partial r_1^2} + 2\epsilon^2 \frac{s}{r_1} \frac{\partial}{\partial r_1} + 2i\epsilon \frac{\omega}{c_R} \frac{\partial}{\partial r_1} + 2\epsilon \frac{is\omega}{r_1 c_R}\right)\widetilde{v}$$

$$= -\frac{2i\omega^3 \mu}{\rho c_R^5 \zeta} \int_{-\infty}^{\infty} \frac{\omega'(\omega-\omega')}{|\omega'(\omega-\omega')|} R(\omega',\omega-\omega')$$

$$\times \widetilde{v}(r_1,\omega')\widetilde{v}(r_1,\omega-\omega') \frac{d\omega'}{2\pi}, \tag{95}$$

where $s = 1/2$ for cylindrical waves, and to recover the corresponding result for plane waves we may let $s = 0$, $r_0 = 0$, and replace r by x. The first two terms on the left-hand side are $O(\epsilon^3)$ and are discarded within the framework of the quadratic approximation. Returning to the space variable r, it can be seen that the progressive wave equation takes the form

$$\frac{\partial \widetilde{v}}{\partial r} + \frac{s}{r}\widetilde{v} = -\frac{\mu\omega^2}{\rho c_R^4 \zeta} \int_{-\infty}^{\infty} \frac{\omega'(\omega-\omega')}{|\omega'(\omega-\omega')|} R(\omega',\omega-\omega')$$

$$\times \widetilde{v}(r,\omega')\widetilde{v}(r,\omega-\omega') \frac{d\omega'}{2\pi}. \tag{96}$$

Equation (96) is analogous to the spectral form of the generalized lossless Burgers equation[35] used in nonlinear acoustics, in that the left-hand sides are the same, and the right-hand sides contain the nonlinear term corresponding to the collinear interaction of plane waves.

We note that different approaches were used to derive the plane progressive wave equation here [Eq. (96) with $s = 0$] and in Sec. I [Eq. (57)]. Equation (96) was obtained via the introduction of a slow scale, assuming a solution corresponding to a forward propagating wave, in order to reduce the second-order equation to one of first order in the spatial derivatives. Equation (57) was obtained via the method of averaging. The discrete form of Eq. (96) is given by Eq. (B8), which is identical to the equation used in Refs. 1 and 29.

F. Diffracting beams

To describe the propagation of a directive beam, diffraction effects are taken into account within the parabolic approximation, as in the derivation of the nonlinear equation for sound beams in fluids.[36,37] The coordinate system is taken such that the x axis coincides with the direction of propagation. The source is located at $x = 0$, is defined as a strip of characteristic width $2a$ along the y axis, and extends to infinite depth within the solid. Radiated frequencies are assumed to be such that $ka \gg 1$, and due to the assumed directivity of the source, the wave is quasiplanar. The retarded time $t' = t - x/c_R$ is introduced, and the Fourier transform of the velocity function in a coordinate system that moves with speed c_R is related to $\widetilde{V}(x,y,\omega)$ by

$$\widetilde{v}(x,y,\omega) = e^{-i\omega x/c_R}\widetilde{V}(x,y,\omega). \tag{97}$$

For a directive beam, with the assumption that effects of diffraction and nonlinearity contribute at the same order to waveform evolution, the appropriate slow scales for the direction of propagation and for the transverse coordinate, respectively, are introduced as

$$x_1 = \epsilon x, \quad y_1 = \epsilon^{1/2} y. \tag{98}$$

After transforming from the coordinates (x,y) to (x_1,y_1) and dropping $O(\epsilon^3)$ terms, one finds that Eq. (93) for $\widetilde{v}(x_1,y_1,\omega)$ becomes

$$\left(\epsilon \frac{\partial^2}{\partial y_1^2} + 2i\epsilon \frac{\omega}{c_R} \frac{\partial}{\partial x_1}\right)\widetilde{v}$$

$$= -\frac{2i\omega^3 \mu}{\rho c_R^5 \zeta} \int_{-\infty}^{\infty} \frac{\omega'(\omega-\omega')}{|\omega'(\omega-\omega')|} R(\omega',\omega-\omega')$$

$$\times \widetilde{v}(x_1,y_1,\omega')\widetilde{v}(x_1,y_1,\omega-\omega') \frac{d\omega'}{2\pi}. \tag{99}$$

In (x,y) coordinates, Eq. (99) may be written

$$\left(\frac{\partial}{\partial x} + \frac{c_R}{2i\omega} \frac{\partial^2}{\partial y^2}\right)\widetilde{v}$$

$$= -\frac{\omega^2 \mu}{\rho c_R^4 \zeta} \int_{-\infty}^{\infty} \frac{\omega'(\omega-\omega')}{|\omega'(\omega-\omega')|} R(\omega',\omega-\omega')$$

$$\times \widetilde{v}(x,y,\omega')\widetilde{v}(x,y,\omega-\omega') \frac{d\omega'}{2\pi}. \tag{100}$$

Equation (100) is equivalent to the result obtained by Kalyanasundaram et al.,[26] and it is analogous to the spectral form of the KZ equation[36,37] for nonlinear sound beams in fluids. The discrete form of Eq. (100) is given by Eq. (B11), and it is identical to the equation solved in Ref. 30 for nonlinear effects in directive Rayleigh wave beams.

III. CONCLUSION

Model equations for nonlinear Rayleigh waves, which are not restricted to periodic waveforms and plane waves, were derived using the techniques developed by Zabolotskaya.[1] The coupled spectral equations used previously to study nonlinear effects in cylindrical waves[1,29] and diffracting beams[30] were derived rigorously.

APPENDIX A: COMPARISON WITH PARKER'S THEORY

Here we demonstrate that the plane-wave theory obtained by Zabolotskaya[1] is equivalent to a theory obtained by Parker.[2] To make this comparison, the appropriate form of the present theory is Eq. (54) for the evolution of a continuous spectrum associated with the particle displacement. The same approximations were used by both authors, i.e., interactions between surface wave and bulk wave modes were neglected, and weak nonlinearity was assumed. Whereas the present theory was derived with Hamiltonian formalism, Parker's theory originates from a dynamical approach. Specifically, the governing nonlinear differential equations and the nonlinear free-surface boundary conditions were used as

a starting point. Parker obtained the evolution equation for the spectrum of the vertical displacement component as a compatibility condition that emerges in the solution of a nonlinear boundary value problem.

In the work of Parker and others, Murnaghan's third-order elastic constants,[38] l, m, and n, are used, whereas our results are expressed in terms of Landau's third-order elastic constants,[39] A, B, and C. We therefore begin by writing the coefficients α', β', and γ' that appear in the kernels $w(\omega_1,\omega_2,\omega_3)$ [Eq. (9)] and $R(\omega_1,\omega_2)$ [Eq. (53)] in terms of the Landau constants:[1]

$$\alpha' = \alpha + \delta, \tag{A1}$$

$$\beta' = \frac{\eta^2 \xi_t}{\mu} (7\mu/3 + A + 2B + K)(1-\xi_l^4), \tag{A2}$$

$$\gamma' = \gamma + \frac{1}{3}\left[\nu + \frac{2\xi_l(\epsilon-2\delta)}{(\xi_t+\xi_l)}\right], \tag{A3}$$

where

$$\alpha = \frac{\eta\xi_t^2}{\mu}(7\mu/3 + A + 2B + K)(1-\xi_l^2), \tag{A4}$$

$$\gamma = \frac{\eta^3(1-\xi_l^2)}{\mu}\left[(\mu+A/3)(1+\xi_l^2+\xi_l^4)\right.$$
$$\left. + (B+K/2-\mu/3)(1+\xi_l^4) + \frac{C}{3}(1-\xi_l^2)^2\right], \tag{A5}$$

$$\delta = \frac{\eta(1-\xi_l^2)}{\mu}\left[(\mu+A/4)(1+\xi_t^2+\xi_t^4) + \frac{A\xi_t^2}{4}\right.$$
$$\left. + \frac{B}{2}(1+\xi_t^2)^2 + (K/2-\mu/3)(1+\xi_t^4)\right], \tag{A6}$$

$$\epsilon = \frac{\eta^2\xi_l}{\mu}(7\mu/3 + A + 2B + K)(1+\xi_t^2)(1-\xi_l^2), \tag{A7}$$

$$\nu = \frac{\eta^3\xi_l^2}{\mu}(7\mu/3 + A + 2B + K)(1-\xi_l^2). \tag{A8}$$

Here, K is the bulk modulus and μ is the shear modulus. Using standard relations[40] between the Landau and Murnaghan constants we obtain the alternative expressions

$$\alpha' = \frac{\eta(1-\xi_l^2)}{4}\left[8\xi_t\nu_1(\xi_t+\xi_l) + \frac{(\lambda+\mu)}{\mu}(1-\xi_t^2)^2\right], \tag{A9}$$

$$\beta' = 2\eta^2\xi_t(1-\xi_l^4)\nu_1, \tag{A10}$$

$$\gamma' = \frac{\eta(1-\xi_l^2)}{3}\left[\left(2\eta^2(1+\xi_l^2)^2 - \frac{16\xi_l^2\xi_t}{(\xi_l+\xi_t)}\right)\nu_1 + \eta^2(1\right.$$
$$\left. -\xi_l^2)^2\nu_2 - \frac{\xi_l}{(\xi_l+\xi_t)}\frac{(\lambda+\mu)}{\mu}(1-\xi_t^2)^2\right], \tag{A11}$$

where λ is the Lamé constant, with ν_1 and ν_2 as defined by Parker:[2]

$$\nu_1 = (\lambda + 3\mu + 2m)/2\mu, \tag{A12}$$

$$\nu_2 = (\lambda + 2l)/2\mu. \tag{A13}$$

Note that Eqs. (A9)–(A13) are independent of n. The reason becomes clear if we examine the expression for the strain energy density in the form used by Murnaghan:[38]

$$\mathscr{E}_2 + \mathscr{E}_3 = \tfrac{1}{2}(\lambda+2\mu)I_1^2 - 2\mu I_2 + \tfrac{1}{3}(l+2m)I_1^3 - 2mI_1I_2$$
$$+ nI_3, \tag{A14}$$

where $I_1 = u_{ii}$, $I_2 = \tfrac{1}{2}(I_1^2 - u_{ij}u_{ij})$, and $I_3 = \det(u_{ij})$ are invariants of the strain tensor:

$$u_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{\partial u_l}{\partial x_i}\frac{\partial u_l}{\partial x_j}\right). \tag{A15}$$

In the case of plane waves, the displacement vector given by Eqs. (1) and (2) has no y component, so $I_3 = \det(u_{ij}) = 0$, and the strain energy density does not depend on n. For general 2-D propagation, the nonlinear coupling coefficients were taken as for collinear interactions, and therefore $I_3 = 0$ as well.

Parker expresses the vertical component of the displacement at the surface in the form

$$U_z(\theta,X) = \int_{-\infty}^{\infty} C(k,X)e^{ik\theta}\,dk, \tag{A16}$$

where the dimensionless variables are defined as follows:

$$U_z = u_z/L \quad \theta = (x - c_R t)/L, \quad X = \epsilon x/L. \tag{A17}$$

The length scale L is used by Parker to designate a characteristic wavelength, with $\epsilon = u_0/L \ll 1$ being the same small quantity used in the present article. The coupled equations obtained by Parker [his Eq. (5.9)] for the variation of the Fourier transform of the vertical displacement have the form

$$\frac{J}{\bar{R}}\frac{\partial C}{\partial X} = i\int_0^k \bar{\alpha}(k_1/k)(k-k_1)k_1 C(k-k_1,X)$$
$$\times C(k_1,X)dk_1 + i\int_0^\infty \bar{\beta}(k_2/k)(k+k_2)$$
$$\times k_2 C(k+k_2,X)C^*(k_2,X)dk_2. \tag{A18}$$

To avoid confusion between Parker's notation and ours, we have used $\bar{R}$, $\bar{\alpha}$, and $\bar{\beta}$ to designate his constant R and functions α and β. The kernels $\bar{\alpha}(k_1/k)$ and $\bar{\beta}(k_2/k)$ are given in Parker's notation by

$$\bar{R}\bar{\alpha}(k_1/k) = -i\left(\frac{G_{111}}{2\beta_1} + \frac{G_{122}+G_{211}}{\beta_1+\beta_2} + \frac{G_{222}}{2\beta_2}\right)$$
$$- \frac{iG_{112}}{\beta_1+\beta_2}\left(\frac{k}{\Gamma(k-k_1)+k_1} + \frac{k}{k-k_1+\Gamma k_1}\right)$$
$$- \frac{iG_{212}}{\beta_1+\beta_2}\left(\frac{k}{k-k_1+(2-\Gamma)k_1}\right.$$
$$\left. + \frac{k}{k_1+(2-\Gamma)(k-k_1)}\right), \tag{A19}$$

$$\bar{R}\bar{\beta}(k_2/k) = -\frac{i2k}{k+k_2}\left(\frac{G_{111}}{2\beta_1} + \frac{G_{122}+G_{211}}{\beta_1+\beta_2} + \frac{G_{222}}{2\beta_2}\right)$$

$$-\frac{iG_{112}}{\beta_1+\beta_2}\left(\frac{2k}{\Gamma k+k_2} + \frac{2k}{k+\Gamma k_2}\right)$$

$$-\frac{iG_{212}}{\beta_1+\beta_2}\left(\frac{2k}{k+(2-\Gamma)k_2}\right.$$

$$\left.+\frac{2k}{(2-\Gamma)k+k_2}\right), \tag{A20}$$

with the following correspondence with the notation used in the present article:

$$\beta_1 = \xi_t, \quad \beta_2 = \xi_l, \quad \Gamma = \frac{2\xi_t}{\xi_t+\xi_l}. \tag{A21}$$

The constants J and $\bar{R}$ appearing in Eqs. (A18)–(A20) are expressed through material constants as follows:

$$J = \frac{4\mu}{1-\beta_1^2}\left(\frac{1-5\beta_1^2+2\beta_1^4}{\beta_1(1+\beta_1^2)} + \frac{\beta_1}{\beta_2^2}\right), \tag{A22}$$

$$\bar{R} = -i\left(\frac{G_{111}}{2\beta_1} + \frac{G_{122}+G_{211}}{\beta_1+\beta_2} + \frac{G_{222}}{2\beta_2}\right), \tag{A23}$$

with

$$G_{111} = 0, \tag{A24}$$

$$G_{211} = \frac{8i\beta_1(1-\beta_2^2)}{1-\beta_1^4}\left[\frac{1}{2}(\lambda+\mu) + \mu\nu_1\right], \tag{A25}$$

$$G_{122} = \frac{64i\beta_1^2}{(1-\beta_1^2)(1-\beta_1^4)^2}(1-\beta_2^2)(\beta_1+\beta_2)$$

$$\times(1+\beta_1\beta_2)\mu\nu_1, \tag{A26}$$

$$G_{222} = -\frac{64i\beta_1^3}{(1-\beta_1^4)^3}(1-\beta_2^2)[2(1+\beta_2^2)^2\mu\nu_1$$

$$+(1-\beta_2^2)^2\mu\nu_2], \tag{A27}$$

$$G_{112} = -\frac{16i\beta_1(1-\beta_2^2)}{(1-\beta_1^2)^2(1-\beta_1^4)}\left[\frac{1}{4}(1-\beta_1^2)^2(\lambda+\mu)\right.$$

$$\left.+2\beta_1(\beta_1+\beta_2)\mu\nu_1\right], \tag{A28}$$

$$G_{212} = \frac{64i\beta_1^3(1-\beta_2^2)}{(1-\beta_1^2)(1-\beta_1^4)^2}\mu\nu_1. \tag{A29}$$

Further manipulations yield

$$J = \frac{4\mu\zeta}{(1-\xi_t^2)}, \tag{A30}$$

$$\bar{R} = \frac{8\mu}{(1-\xi_t^2)^3}\left(\frac{3\gamma'}{2\xi_l} + \frac{\alpha'+\beta'}{\xi_l+\xi_t}\right), \tag{A31}$$

$$G_{112} = \frac{8i\mu}{(1-\xi_t^2)^3}\alpha', \tag{A32}$$

$$G_{212} = \frac{8i\mu}{(1-\xi_t^2)^3}\beta'. \tag{A33}$$

Expressing Parker's coefficients in terms of the notation used in the present article required extensive use of the dispersion relation for a Rayleigh wave,

$$(1+\xi_t^2)^2 = 4\xi_t\xi_l. \tag{A34}$$

Finally, substitution of the above coefficients into Eqs. (A19) and (A20) yields

$$\bar{R}\bar{\alpha}(k_1/k) = \frac{8\mu|k|}{(1-\xi_t^2)^3}R(k_1,k-k_1), \tag{A35}$$

$$\bar{R}\bar{\beta}(k_2/k) = \frac{16\mu|k|}{(1-\xi_t^2)^3}R(k_2+k,-k_2). \tag{A36}$$

After substituting Eqs. (A35) and (A36) into Eq. (A18) and expressing $C(k,X)$ from Ref. 2 in terms of $\tilde{u}(x,\omega)$ [see Eqs. (A16) and (42) for the vertical displacement component], our Eq. (54) for $\tilde{u}(x,\omega)$ is recovered exactly. The model equations derived by Zabolotskaya[1] and Parker[2] are thus equivalent.

APPENDIX B: DISCRETE FORMS OF THE SPECTRAL EQUATIONS

To rewrite the equations derived in this article in discrete form for a periodic wave, which is the form most amenable to numerical solution, the velocity components are expressed as Fourier series:

$$\mathbf{v}(\mathbf{r},z,t) = \frac{1}{2}\sum_{n=-\infty}^{\infty}\frac{i}{nk_r}h_\perp(z,n\omega_r/c_R)\boldsymbol{\nabla}_\perp\tilde{V}_n(\mathbf{r})e^{-in\omega_r t}, \tag{B1}$$

$$v_z(\mathbf{r},z,t) = \frac{1}{2}\sum_{n=\infty}^{\infty}h_z(z,n\omega_r/c_R)\tilde{V}_n(\mathbf{r})e^{-in\omega_r t}. \tag{B2}$$

Here $k_r = \omega_r/c_R$, and ω_r is the fundamental, or repetition, frequency. The depth-dependent functions $h_\perp(z,n\omega_r/c_R)$ and $h_z(z,n\omega_r/c_R)$ are given by Eqs. (60) and (61) with $|\mathbf{k}| = nk_r$. Comparing Eqs. (B1) and (B2) with the continuous spectrum representations of $\mathbf{v}(\mathbf{r},z,t)$ and $v_z(\mathbf{r},z,t)$ given by Eqs. (89) and (90), we make the substitutions

$$\tilde{V}(\mathbf{r},\omega=n\omega_r)\to\tilde{V}_n(\mathbf{r})/\omega_r, \tag{B3}$$

$$\int_{-\infty}^{\infty}\frac{d\omega_1}{2\pi\omega_r}\to\frac{1}{2}\sum_{m=-\infty}^{\infty}, \tag{B4}$$

to rewrite Eq. (93) for a periodic signal as

$$\nabla_\perp^2\tilde{V}_n + \frac{n^2\omega_r^2}{c_R^2}\tilde{V}_n = i\frac{\mu\omega_r^2 n^3}{\rho c_R^5\zeta}\left(2\sum_{m=n+1}^{\infty}R_{m,n-m}\tilde{V}_m\tilde{V}_{m-n}^*\right.$$

$$\left.-\sum_{m=1}^{n-1}R_{m,n-m}\tilde{V}_m\tilde{V}_{n-m}\right), \tag{B5}$$

where $R_{m,n-m}\equiv R(m,n-m)$ [recall Eq. (53)]. Note that Eq. (B5) describes the variation of a complex spectral amplitude V_n with $n>0$. Velocity amplitudes with $n<0$ are obtained from the relation $\tilde{V}_{-n}=\tilde{V}_n^*$, which follows from the require-

1415   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Knight *et al.*: Nonlinear Rayleigh wave theory   1415

ment that the velocities be real. For a finite Fourier series approximation of the spectrum, $1 \leqslant n \leqslant N$, replace the infinite limit on the first summation in Eq. (B5) with N.

To obtain the discrete form of the model equation for a progressive one-dimensional wave, the velocity components are represented as

$$v_r(r,z,t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} i \frac{n}{|n|} h_\perp(z, n\omega_r/c_R)$$
$$\times \widetilde{v}_n(r) e^{in[k_r(r-r_0)-\omega_r t]}, \tag{B6}$$

$$v_z(r,z,t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} h_z(z, n\omega_r/c_R)$$
$$\times \widetilde{v}_n(r) e^{in[k_r(r-r_0)-\omega_r t]}, \tag{B7}$$

where r is replaced by x, with $r_0 = 0$, for a plane wave. Substitution of Eqs. (B3) and (B4) into Eq. (96) for one-dimensional waves yields

$$\frac{d\widetilde{v}_n}{dr} + \frac{s}{r} \widetilde{v}_n = \frac{\mu \omega_r n^2}{2\rho c_R^4 \zeta} \left(2 \sum_{m=n+1}^{\infty} R_{m,n-m} \widetilde{v}_m \widetilde{v}_{m-n}^* \right.$$
$$\left. - \sum_{m=1}^{n-1} R_{m,n-m} \widetilde{v}_m \widetilde{v}_{n-m} \right), \tag{B8}$$

with $s = 0$ for a plane wave and $s = 1/2$ for a cylindrical wave.

To obtain the discrete form of Eq. (100) for a diffracting beam, the velocity components should be expressed in the form

$$v_x(x,y,z,t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} i \frac{n}{|n|} h_\perp(z, n\omega_r/c_R)$$
$$\times \widetilde{v}_n(x,y) e^{in(k_r x - \omega_r t)}, \tag{B9}$$

$$v_z(x,y,z,t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} h_z(z, n\omega_r/c_R)$$
$$\times \widetilde{v}_n(x,y) e^{in(k_r x - \omega_r t)}. \tag{B10}$$

Substituting Eqs. (B3) and (B4) into Eq. (100), we obtain

$$\frac{\partial \widetilde{v}_n}{\partial x} + \frac{1}{2ink_r} \frac{\partial^2 \widetilde{v}_n}{\partial y^2} = \frac{\mu \omega_r n^2}{2\rho c_R^4 \zeta} \left(2 \sum_{m=n+1}^{\infty} R_{m,n-m} \widetilde{v}_m \widetilde{v}_{m-n}^* \right.$$
$$\left. - \sum_{m=1}^{n-1} R_{m,n-m} \widetilde{v}_m \widetilde{v}_{n-m} \right). \tag{B11}$$

[1] E. A. Zabolotskaya, ''Nonlinear propagation of plane and circular Rayleigh waves in isotropic solids,'' J. Acoust. Soc. Am. **91**, 2569–2575 (1992).

[2] D. F. Parker, ''Waveform evolution for nonlinear surface acoustic waves,'' Int. J. Eng. Sci. **26**, 59–75 (1988).

[3] J. W. S. Rayleigh, ''On waves propagated along the plane surface of an elastic solid,'' Proc. London Math. Soc. **17**, 4–11 (1885).

[4] I. A. Viktorov, ''Effects of a second approximation in the propagation of waves through solids,'' Sov. Phys. Acoust. **9**, 242–245 (1964).

[5] F. Rischbieter, ''Messungen an Oberflächenwellen in festen Körpern,'' Acustica **16**, 75–83 (1965–1966).

[6] F. Rischbieter, ''Messung des nichtlinearen Schallverhaltens von Aluminium mit Hilfe von Rayleighwellen,'' Acustica **18**, 109–112 (1967).

[7] D. P. Morgan, *Surface-Wave Devices for Signal Processing* (Elsevier, New York, 1985).

[8] A. A. Kolomeskii and A. A. Maznev, ''Shake-off of mechanical microparticles from a silicon surface with surface acoustic waves induced by laser pulses,'' Sov. Tech. Phys. Lett. **17**, 483–484 (1991).

[9] E. Yu. Knight, ''Generalization of the theory for nonlinear Rayleigh waves to nonplanar and transient waveforms, and investigation of pulse propagation,'' M.A. thesis, The University of Texas at Austin, 1995.

[10] E. G. Lean and C. C. Tseng, ''Nonlinear effects in surface acoustic waves,'' J. Appl. Phys. **41**, 3912–3917 (1970).

[11] E. L. Adler, E. Bridoux, G. Coussot, and E. Dieulesaint, ''Harmonic generation of acoustic surface waves in Bi$_{12}$GeO$_{20}$ and LiNbO$_3$,'' IEEE Trans. Sonics Ultrason. **SU-20**, 13–16 (1973).

[12] N. Kalyanasundaram, ''Nonlinear surface acoustic waves on an isotropic solid,'' Int. J. Eng. Sci. **19**, 279–286 (1981).

[13] P. O. Løpen, ''Second harmonic generation in an elastic surface wave in α-quartz,'' J. Appl. Phys. **39**, 5400–5404 (1968).

[14] V. E. Ljamov, T. H. Hsu, and R. M. White, ''Surface elastic wave velocity and second harmonic generation in an elastically nonlinear medium,'' J. Appl. Phys. **43**, 800–804 (1972).

[15] H. F. Tiersten and J. C. Baumhauer, ''Second harmonic generation and parametric excitation of surface waves in elastic and piezoelectric solids,'' J. Appl. Phys. **45**, 4272–4287 (1974).

[16] P. J. Vella, T. C. Padmore, G. I. Stegeman, and V. M. Ristic, ''Nonlinear surface-wave interactions: Parametric mixing and harmonic generation,'' J. Appl. Phys. **45**, 1993–2006 (1974).

[17] N. Kalyanasundaram, R. Ravindran, and P. Prasad, ''Coupled amplitude theory of nonlinear surface acoustic waves,'' J. Acoust. Soc. Am. **72**, 488–493 (1982).

[18] N. Kalyanasundaram and G. V. Anand, ''Periodic Rayleigh waves of finite amplitude on an isotropic solid,'' J. Acoust. Soc. Am. **72**, 1518–1523 (1982).

[19] N. Kalyanasundaram, ''Nonlinear mixing of surface acoustic waves propagating in opposite directions,'' J. Acoust. Soc. Am. **73**, 1956–1965 (1983).

[20] R. W. Lardner, ''Nonlinear surface waves on an elastic solid,'' Int. J. Eng. Sci. **21**, 1331–1342 (1983).

[21] R. W. Lardner, ''Nonlinear Rayleigh waves: Harmonic generation, parametric amplification, and thermoviscous damping,'' J. Appl. Phys. **55**, 3251–3260 (1984).

[22] R. W. Lardner, ''Waveform distortion and shock development in nonlinear Rayleigh waves,'' Int. J. Eng. Sci. **23**, 113–118 (1985).

[23] E. A. David, ''A uniform asymptotic solution for nonlinear surface acoustic waves,'' Int. J. Eng. Sci. **23**, 699–708 (1985).

[24] D. F. Parker and F. M. Talbot, ''Analysis and computation for nonlinear elastic surface waves of permanent form,'' J. Elast. **15**, 389–426 (1985).

[25] F. John, ''Plane strain problems for a perfectly elastic material of harmonic type,'' Commun. Pure Appl. Math. **13**, 239–296 (1960).

[26] N. Kalyanasundaram, D. F. Parker, and E. A. David, ''The spreading of nonlinear elastic surface waves,'' J. Elast. **24**, 79–103 (1990).

[27] Y. Shui and I. Yu. Solodov, ''Nonlinear properties of Rayleigh and Stoneley waves in solids,'' J. Appl. Phys. **64**, 6155–6165 (1988).

[28] M. F. Hamilton, Yu. A. Il'insky, and E. A. Zabolotskaya, ''On the existence of stationary nonlinear Rayleigh waves,'' J. Acoust. Soc. Am. **93**, 3089–3095 (1993).

[29] D. J. Shull, M. F. Hamilton, Yu. A. Il'insky, and E. A. Zabolotskaya, ''Harmonic generation in plane and cylindrical nonlinear Rayleigh waves,'' J. Acoust. Soc. Am. **94**, 418–427 (1993).

[30] D. J. Shull, E. E. Kim, M. F. Hamilton, and E. A. Zabolotskaya, ''Diffraction effects in nonlinear Rayleigh wave beams,'' J. Acoust. Soc. Am. **97**, 2126–2137 (1995).

[31] M. F. Hamilton, Yu. A. Il'insky, and E. A. Zabolotskaya, ''Local and nonlocal nonlinearity in Rayleigh waves,'' J. Acoust. Soc. Am. **97**, 882–890 (1995).

[32] M. F. Hamilton, Yu. A. Il'insky, and E. A. Zabolotskaya, ''Evolution equations for nonlinear Rayleigh waves,'' J. Acoust. Soc. Am. **97**, 891–897 (1995).

[33] N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-Linear Oscillations* (Hindustan, India, 1961).

[34] P. J. Westervelt, ''Parametric acoustic array,'' J. Acoust. Soc. Am. **35**, 535–537 (1963).

[35] D. T. Blackstock, ''On plane, spherical, and cylindrical sound waves of finite amplitude in lossless fluids,'' J. Acoust. Soc. Am. **36**, 217–219 (1964).

1416   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Knight *et al.*: Nonlinear Rayleigh wave theory   1416

[36] E. A. Zabolotskaya and R. V. Khokhlov, ''Quasi-plane waves in the nonlinear acoustics of confined beams,'' Sov. Phys. Acoust. **15**, 35–40 (1969).

[37] N. S. Bakhvalov, Ya. M. Zhileikin, and E. A. Zabolotskaya, *Nonlinear Theory of Sound Beams* (American Institute of Physics, New York, 1987).

[38] F. Murnaghan, *Finite Deformation of an Elastic Solid* (Dover, New York, 1951).

[39] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, New York, 1986), 3rd ed., p. 107.

[40] R. E. Green, *Ultrasonic Investigation of Mechanical Properties* (Academic, New York, 1973).

The effects of surfactant additives on the acoustic and light emissions from a single stable sonoluminescing bubble

Thomas R. Stottlemyer
Naval Undersea Warfare Center, Code 3113, Newport, Rhode Island 02841

Robert E. Apfel
Department of Mechanical Engineering, Yale University, New Haven, Connecticut 06520

The localized concentration of energy during a single bubble collapse is manifested in two forms, which are the emission of an acoustic pulse, and the emission of a light pulse. Through precise control of experimental parameters, one can levitate a single bubble in a standing wave field and measure the magnitude of the acoustic and light pulses resulting from the violent collapse of the cavity. The information acquired from such measurements provides better understanding of the mechanisms that are responsible for the emissions, which may lead to the practical application of controlled bubble implosions. An experimental apparatus was developed to measure the acoustic and light emissions from a single, stable sonoluminescing bubble. Two surfactant additives were studied to determine the effects on the bubble emissions. Triton X-100, which has previously been shown to provide free interfacial motion, reduced the magnitude of both the acoustic and light pulses from the bubble. The protein bovine serum albumin (BSA) which has been shown to hinder interfacial motion, allowed the bubble to be driven to higher acoustic pressures, and resulted in an increase in the magnitude of the light pulses from the bubble. Images of the sonoluminescing bubble indicate that the collapse remains spherical in the cases presented, and that bubble translation can be correlated with weak acoustic and light emissions. © *1997 Acoustical Society of America.* [S0001-4966(97)02809-9]

PACS numbers: 43.25.Ts, 43.35.Sx [MAB]

INTRODUCTION

The collapse of a spherical cavity in a liquid has been a topic of considerable interest since the theoretical investigations of Rayleigh.[1] More recent research by Gaitan *et al.*[2] led to the discovery that a single bubble could be levitated in a standing wave field, and the acoustic pressure increased to a point at which light pulses, or sonoluminescence, synchronized to the acoustic cycle, resulted from the repeated violent collapse of the cavity. Light scattering techniques[2,3] have been used to measure the radius of the sonoluminescing bubble versus time, and changes in experimental parameters have been investigated to determine the effects on the intensity of the light pulses from the bubble.[4,5] Cordry[6] developed a technique to measure the acoustic pulses from a sonoluminescing bubble by placing a needle hydrophone adjacent to the bubble, providing the ability to study the effects of fluid temperature changes on the acoustic output from the bubble. Each of these investigations has led to a better understanding of the characteristics of the light pulse that is generated upon bubble collapse, and of the mechanisms whereby light is generated.

Stebe and Maldarelli[7] determined, by conducting capillary slug flow experiments, that the mobility of interfaces coated with surfactants were altered by varying the type of surfactant, and the concentration about the critical micelle concentration (CMC).

Studies of the magnitude of the acoustic pulses and the light pulses from a single stable sonoluminescing bubble are presented here and yield information about the behavior of the bubble in the presence of surfactant additives.[8] Triton X-100, which has been shown to provide free interfacial motion, and the protein bovine serum albumin (BSA), which hinders surface mobility,[7] were chosen for the experiments that were conducted.

I. EXPERIMENTAL APPARATUS

The experimental apparatus that was used to levitate a single stable sonoluminescing bubble and to measure the acoustic and light emissions from the bubble is shown in Fig. 1. The test cell consists of Plexiglas plates joined to form a container with dimensions of 6.91 cm×6.91 cm×15.24 cm. Attached to the bottom of the test cell is a cylindrical lead zirconate titanate (PZT) transducer (EDO Corporation Western Division) with 3.81 cm outside diameter and height, and 3.18 mm thickness. A needle hydrophone with a length of 25.5 cm and a diameter of 1.27 mm was mounted above the test cell, and lowered into the cell, adjacent to the bubble. The needle hydrophone provided the ability to measure the relative magnitude of the acoustic pulses from the sonoluminescing bubble along with the acoustic driving pressure in the test cell. A Gaertner filar microscope was mounted outside the cell and was used for precise placement of the bubble relative to the needle hydrophone. An EG&G ORTEC Model 9201 Photomultiplier Tube (PMT) was located opposite the microscope and was used to measure the relative magnitude of the light pulses from the bubble. Signals from the PMT and the needle hydrophone were acquired

II. EXPERIMENTAL PROCEDURE

The procedure for acquiring the acoustic and light pulses from the sonoluminescing bubble consisted of first aligning the needle hydrophone and PMT relative to the bubble. This was required for consistent measurements, and was accomplished by levitating a bubble in the test cell, measuring the distance from the bubble to the lens of the PMT, and using the Gaertner microscope to place the bubble near the needle hydrophone. The distance between the needle hydrophone and the bubble was 1.92 mm for the experiments presented here. The position of the needle hydrophone had to be readjusted frequently because the bubble translated gradually with time, and shifted when the acoustic driving pressure in the test cell was changed. Once the alignment was completed, the signal from the hydrophone was stored on the oscilloscope. The acoustic driving pressure was increased, and the signal was stored. This process was repeated until the extinction threshold of the bubble, or the pressure at which the bubble ceased to exist, was reached. The driving pressure would then be lowered to the initial value, the lights in the ''dark room'' would be turned off, a new bubble would be generated, and the magnitude of the light pulses from the PMT would be measured on the oscilloscope, for the same range of acoustic driving pressures.

In order to measure the magnitude of the acoustic pulses acquired on the oscilloscope, the data were transferred to an Apple Power Macintosh 7100/66 and processed using LabVIEW® (National Instruments Corp.). The program that was written digitally filtered out the driving signal from the oscilloscope trace, leaving the bubble pulses received on the needle hydrophone. The maxima and minima of each pulse was measured, and then the magnitudes of each pulse were averaged over 12 acoustic cycles. It was determined that there was a variation of as much as 15% in the magnitude of the acoustic pulse from one acoustic cycle to the next, which necessitated the averaging routine.

After the acoustic and light pulses had been measured, images of the sonoluminescing bubble were acquired using a strobing technique that was developed by Tian et al.[9] By flashing a light emitting diode (LED) at a frequency just slightly off of the test cell driving frequency, images of the collapse sequence were captured over successive acoustic cycles with a microscope and a video camera. The images were a tool for determining the bubble shape at larger radii, but could not be used to resolve the bubble shape below a radius of about 5 μm. The images were also useful for generating radius versus time curves for the bubbles. The system was calibrated by placing a glass slide, with lines etched at 50-μm spacing, into the water-filled test cell, and acquiring video images. The images of the glass slides provided a calibration in pixels per μm which was used to determine the radii of the bubbles that were acquired with the imaging system. All of the images presented here were acquired at the maximum acoustic driving pressure prior to bubble extinction.

III. RESULTS

In order to provide a basis for comparison, the fluid that was initially investigated was deionized water at room tem-

FIG. 1. Diagram of test cell and electronic equipment.

on a Lecroy 7200 Precision Digital Oscilloscope. The continuous wave (CW) signal to the cylindrical transducer was generated with a Hewlett-Packard model 8904A Multifunction Synthesizer. An ENI Model 1040L Power Amplifier provided signal gain, and an ENI Piezoelectric Transducer Matchbox was used to match the impedance of the transducer to that of the power amplifier. A Hewlett-Packard Model 400E AC Voltmeter was placed between the matchbox and the transducer to monitor the final voltage to the transducer. A probe of the sound field with the needle hydrophone indicated that the test cell was driven in the (1,1,1) mode at a frequency of 16.2 kHz.

The bubble was generated by either injecting air with a small syringe, or by locally boiling the water, thereby creating a vapor cavity, with a NiCr wire. The acoustic pulses and light pulses from sonoluminescing bubble were compared for the two generation techniques, and it was determined that the light and acoustic output from the bubble was a function of the liquid preparation and type, regardless of the method that was used to create the bubble.

Each of the fluids that were used in these experiments were prepared identically. The procedure consisted of placing deionized water in a flask, adding the appropriate amount of surfactant, and degasing the mixture with a vacuum pump. At the completion of the degasing procedure, the gas content in the liquid would be on the order of 1%, which was calculated by using a gas handling system to determine the vapor pressure above the mixture at equilibrium. The fluid mixture was left overnight to equilibrate to room temperature, and then gently poured into the test cell for the measurements. Although it is expected that some regasing during the experiment occurred, the conditions were repeated for all measurements, making relative assessments valid.

FIG. 2. The magnitude of the acoustic pulses from a sonoluminescing bubble in water compared to a solution of 0.1 CMC Triton X-100 in water.

perature. The acoustic and light emissions from the sonoluminescing bubble in water are, in the results that follow, plotted along with the data that was acquired using the surfactant solutions as the host fluid.

Figure 2 provides a relative comparison of the magnitude of the acoustic pulses from a single stable sonoluminescing bubble in deionized water, to a solution comprised of 0.1 CMC of Triton X-100 in deionized water. The figure indicates a reduction in the acoustic output from the bubble, and a negligible difference in the maximum acoustic driving pressure that could be achieved prior to bubble extinction. A comparison of the magnitude of the light pulses is shown in Fig. 3, with a similar reduction in the maximum PMT output. The test cell was driven at a frequency of 16.8 kHz for the experiments using water, and at 16.4 kHz when using the Triton mixture. Experiments were also conducted using concentrations of 0.25 CMC, and 1.0 CMC of Triton X-100 in deionized water. As the concentration of Triton was increased, the bubble exhibited erratic translations in the test cell over the same range of acoustic pressures that provided a stable bubble in the 0.1 CMC Triton solution. The translation of the bubble in the 0.25 CMC solution correlated with very
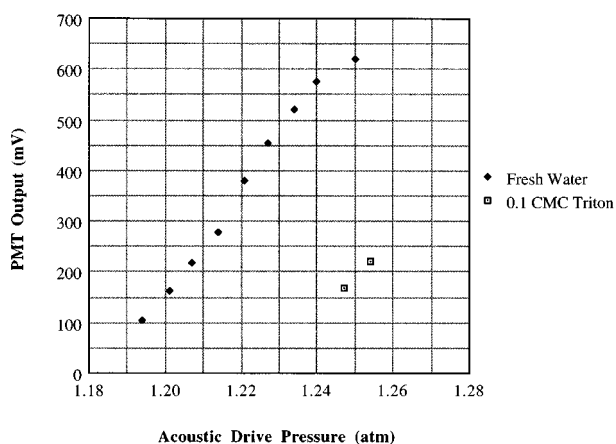
FIG. 3. The magnitude of the light pulses from a sonoluminescing bubble in water compared to a solution of 0.1 CMC Triton X-100 in water.
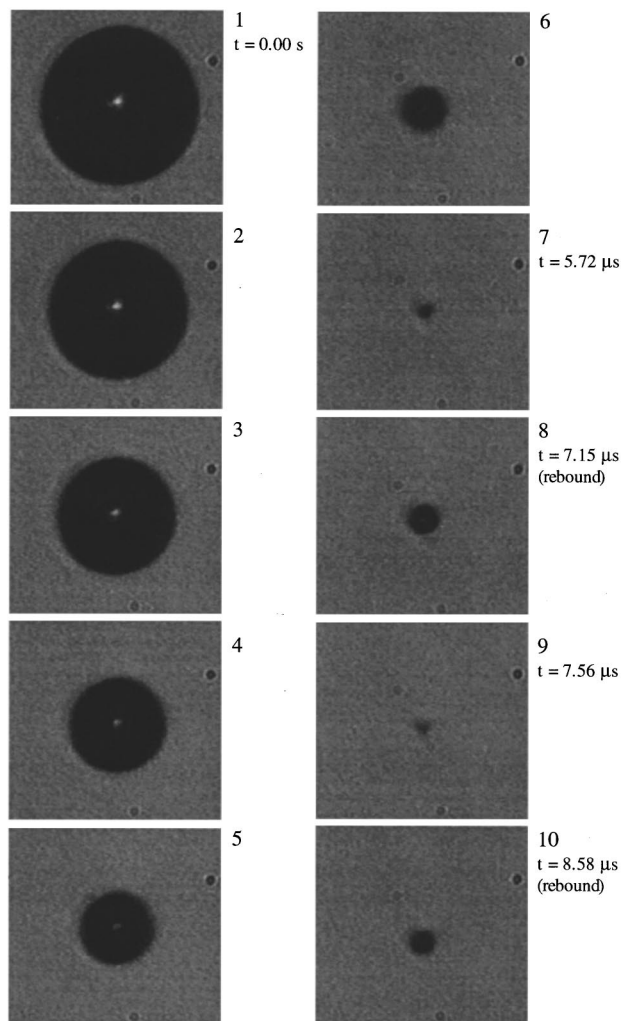
FIG. 4. The collapse of a sonoluminescing bubble in a solution of 0.1 CMC Triton X-100 in water.

weak acoustic pulses, and no detectable light being emitted. The bubble translated even more violently in the 1.0 CMC solution and provided no acoustic or light pulses.

Figure 4 presents images of the sonoluminescing bubble in the 0.1 CMC Triton solution that were acquired using the strobing technique. Only two of the multiple bubble rebounds that occur after the collapse are shown in Fig. 4. The maximum radius is 62 μm, and the minimum radius that is measured with this technique is 9 μm prior to the first rebound (Fig. 5), although it is known that the bubble achieves a much smaller radius at the end of this initial collapse. These images were acquired at the maximum acoustic driving pressure prior to bubble extinction. The maximum radius of the sonoluminescing bubble in deionized water was 65 μm (Fig. 6) which was determined by capturing a collapse sequence that was similar to the one shown for the Triton solution. We speculated that the reduced acoustic and light emissions from the bubble in the Triton mixture may have been a result of a nonspherical collapse, but Fig. 4 indicates that the bubble remains spherical during the entire collapse, and over successive acoustic cycles. Images of the bubble in the 0.25 CMC Triton solution show significant translations, a
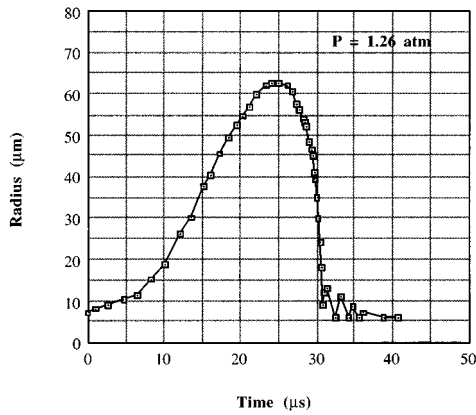
FIG. 5. The radius versus time curve for a sonoluminescing bubble in a solution of 0.1 CMC Triton X-100 in water.

FIG. 7. The magnitude of the acoustic pulses from a sonoluminescing bubble in water compared to solutions of BSA in water.

maximum radius of only 30 μm, and a spherical shape maintained throughout the collapse and rebounds.

A comparison of the magnitude of the acoustic pulses in deionized water to solutions of BSA in deionized water are shown in Fig. 7. The 0.5 CMC BSA solution showed a reduction in the overall magnitude, while the 1.0 CMC BSA solution showed a very slight increase over that of the water. The deionized water and 1.0 CMC BSA maxima are considered statistically equivalent given that the maximum error in the acoustic measurements was determined to be ±0.003 mV. The 0.5 CMC BSA measurements were conducted at a test cell driving frequency of 16.4 kHz, and the 1.0 CMC BSA solution was studied at a frequency of 16.3 kHz. An interesting feature shown in Fig. 7 is the increase in the acoustic driving pressures over which the BSA measurements were conducted. The sonoluminescing bubble was driven to 1.25 atm in water, and to nearly 1.33 atm in the 1.0 CMC BSA solution. In fact, the upper threshold in the BSA solutions was not an extinction threshold as it was for the deionized water and Triton X-100 mixtures, but rather was a stability threshold at which the bubble began to translate. This translation consisted of the bubble oscillating in the vertical direction, while continuing to emit acoustic and light pulses. The initial frequency of the vertical oscillation was approximately 1 Hz, and it increased as the acoustic driving pressure was increased. Measurements of the acoustic pulses
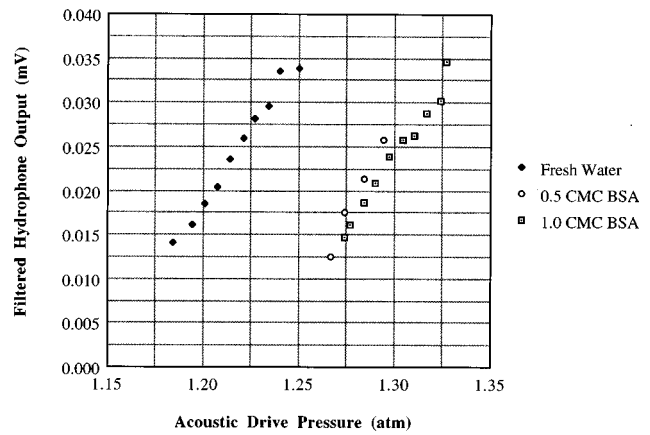
were not conducted above this stability threshold because changes in the location of the bubble relative to the needle hydrophone rendered the magnitude of the acoustic pulses meaningless. As shown in Fig. 7, the bubble was driven to a level of 1.29 atm in the 0.5 CMC BSA mixture, where it began to translate, but the bubble could be driven to an acoustic pressure of 1.39 atm before it suddenly disappeared.

A comparison of the magnitude of the light pulses from bubbles in the different solutions is shown in Fig. 8. The 0.5 CMC BSA solution resulted in a slight increase in the light output from the bubble, while the 1.0 CMC BSA solution resulted in a substantial increase in the light emissions. The maximum error in the light measurements was determined to be ±40 mV.

Images of the sonoluminescing bubble in the 1.0 CMC BSA solution are presented in Fig. 9. The maximum radius of this bubble is 69 μm (Fig. 10) as compared to 65 μm in deionized water, and 62 μm in the 0.1 CMC Triton X-100 solution. As in Fig. 4, the collapse of the bubble in the BSA solution remains spherical during the collapse, and over multiple acoustic cycles.

A drawback of this imaging technique is that the high intensity light required for the magnification prevents the use

FIG. 6. The radius versus time curve for a sonoluminescing bubble in water.
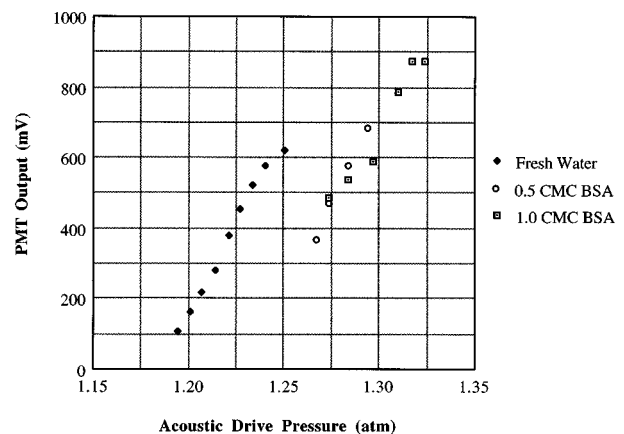
FIG. 8. The magnitude of the light pulses from a sonoluminescing bubble in water compared to solutions of BSA in water.

FIG. 9. The collapse of a sonoluminescing bubble in a solution of 1.0 CMC BSA in water.
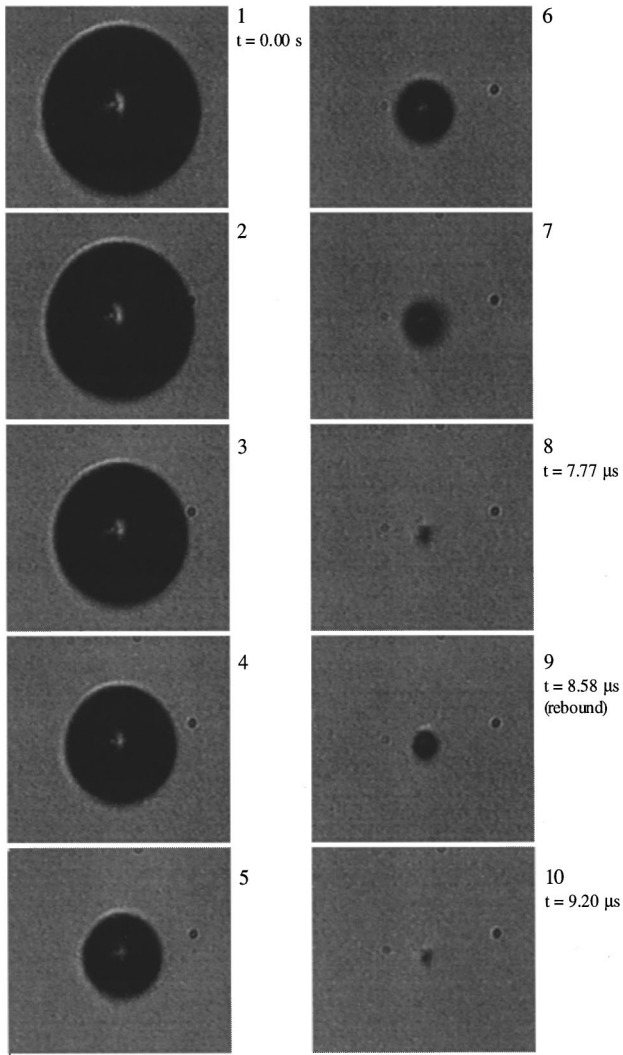
FIG. 10. The radius versus time curve for a sonoluminescing bubble in a solution of 1.0 CMC BSA in water.

of the needle hydrophone near the bubble, because the hydrophone interferes with the LED. Since acoustic driving pressure is not measured directly during the imaging, the acoustic driving pressure is inferred by using the same drive levels of the electronic system as were used during the acoustic and light emission measurements. Therefore, when calculating bubble equilibrium radii using the Rayleigh–Plesset (RP) equation[10] and the van der Waals equation of state, the magnitude of the acoustic driving pressure is not precisely known for the given radius versus time curve. The calculated equilibrium radius is extremely sensitive to the magnitude of the acoustic driving pressure, which was noticed in the case of the sonoluminescing bubble in deionized water. By reducing the acoustic driving pressure by 1% in the calculation, a four-fold increase was observed in the bubble equilibrium radius. Such sensitivity indicates that the calculation of sonoluminescing bubble equilibrium radii will require extremely precise hydrophone calibration. For the three fluids presented here, equilibrium radii of 3 μm were assumed and the calculated acoustic driving pressures[3] agreed with measured values, but without the precise acous-

tic driving pressures associated with the radius versus time curves, the effect of the different fluid types on the bubble equilibrium radius could not be adequately evaluated.
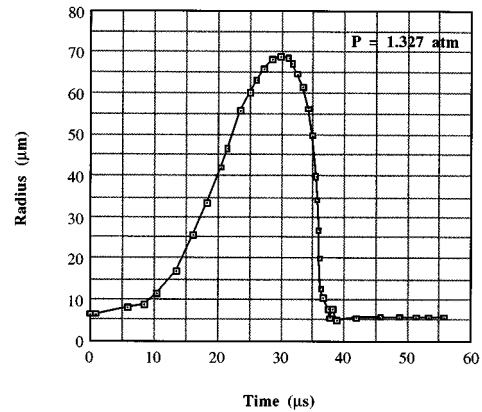
IV. DISCUSSION

The information provided here has increased our understanding of bubble behavior during sonoluminescence, but as questions are answered, nearly as many are developed as a result of the research. A number of trends are apparent from the data presented. Triton X-100 reduced the maximum radius of the bubble prior to collapse, compared to deionized water, thereby decreasing the magnitude of the acoustic and light pulses emitted. As the concentration of Triton X-100 was increased, the bubble began to translate, and this unstable behavior became extremely violent at a concentration of 1.0 CMC. BSA had the effect of increasing the extinction threshold and the maximum radius of the cavity, compared to water, which increased the magnitude of the light pulses substantially.

The effects of the surfactant additives may in part be due to changes in mass diffusion into and out of the bubble. Asaki et al.[11] developed a technique for depositing a known quantity of surfactant on the surface of a radially pulsating bubble, and suggested that the rate of gas diffusion is affected by this layer, which in the case of sonoluminescence could result in a strong effect. The smaller radius in the Triton solution and the larger radius in the BSA solution are quite possibly due to mass diffusion changes.

An additional possibility for the changes in the acoustic and light output from the bubble is related to the surface mobility factors suggested by Stebe and Maldarelli.[7] Since BSA has been shown to hinder the mobility of interfaces, high-frequency capillary waves that develop on the surface of the sonoluminescing bubble may be dampened, resulting in the ability to achieve higher acoustic drive pressures. Triton X-100 which provides freer interfacial mobility had the opposite effect, causing the bubble to translate over a range of acoustic pressures that had previously provided a stable bubble in water, and in solutions with BSA. Brenner et al.[12] suggest that increases in viscosity increase the parametric

instability threshold, thereby causing bubbles in highly viscous liquids to encounter the Rayleigh–Taylor instability prior to the parametric instability. The BSA additive may have altered the interfacial viscosity enough that the parametric instability threshold was increased, without a significant change in the Rayleigh–Taylor instability threshold.

Previous researchers have shown that the behavior of a single stable sonoluminescing bubble is affected by pressure, temperature, the type of gas in the liquid, and the concentration of the gas that is dissolved. The work presented here has shown that in addition to the parameters previously investigated, surfactant additives can affect the acoustic and light pulses from a sonoluminescing bubble quite substantially. We have advanced our knowledge of the behavior of a collapsing cavity in a liquid, which may lead to further use of implosions as controlled phenomena with acoustic or electromagnetic applications.

[1] Lord Rayleigh, ''On the pressure developed in a liquid during the collapse of a spherical cavity,'' Philos. Mag. **34**, 94 (1917).

[2] D. F. Gaitan, L. A. Crum, C. C. Church, and R. A. Roy, ''Sonoluminescence and bubble dynamics for a single, stable, cavitation bubble,'' J. Acoust. Soc. Am. **91**, 3166 (1992).

[3] B. P. Barber and S. J. Putterman, ''Light scattering measurements of the repetitive supersonic implosion of a sonoluminescing bubble,'' Phys. Rev. Lett. **69**, 3839 (1992).

[4] B. P. Barber, C. C. Wu, R. Lofstedt, P. H. Roberts, and S. J. Putterman, ''Sensitivity of sonoluminescence to experimental parameters,'' Phys. Rev. Lett. **72**, 1380 (1994).

[5] B. P. Barber, K. Weninger, R. Lofstedt, and S. J. Putterman, ''Observation of a new phase of sonoluminescence at low partial pressures,'' Phys. Rev. Lett. **74**, 5276 (1995).

[6] S. M. Cordry, ''Bjerknes forces and temperature effects in single-bubble sonoluminescence,'' Ph.D. Dissertation, University of Mississippi (1995).

[7] K. J. Stebe and C. Maldarelli, ''Remobilizing surfactant retarded fluid particle interfaces,'' J. Colloid Interface Sci. **163**, 177 (1994).

[8] T. R. Stottlemyer, ''An experimental study of the acoustic emissions from collapsing cavities in liquids,'' Ph.D. Dissertation, Yale University (1996).

[9] Y. Tian, J. A. Ketterling, and R. E. Apfel, ''Direct observation of microbubble oscillations,'' J. Acoust. Soc. Am. **100**, 3976 (1996).

[10] R. Lofstedt, K. Weninger, S. Putterman, and B. P. Barber, ''Sonoluminescing bubbles and mass diffusion,'' Phys. Rev. E **51**, 4400 (1995).

[11] T. J. Asaki, D. B. Thiessen, and P. L. Marston, ''Effect of an insoluble surfactant on capillary oscillations of bubbles in water: Observation of a maximum in the damping,'' Phys. Rev. Lett. **75**, 2686 (1995).

[12] M. P. Brenner, D. Lohse, and T. F. Dupont, ''Bubble shape oscillations and the onset of sonoluminescence,'' Phys. Rev. Lett. **75**, 954 (1995).

Radiation forces between two bubbles in a compressible liquid

A. A. Doinikov and S. T. Zavtrak

Institute of Nuclear Problems, Bobruiskaya 11, 220050 Minsk, Belarus

The expression for the interaction force exerted by a sound field between two gas bubbles, allowing for the compressibility of the surrounding liquid, is derived. The bubble radii are considered to be small compared to the distance l between the equilibrium centers of the bubbles and the wavelength λ of sound, while the ratio of l to λ is assumed to be arbitrary. For an incompressible liquid, the interaction force, usually called the mutual or secondary Bjerknes force, is known to be inversely proportional to l^2. It is found that when it is taken into account, the compressibility of the liquid gives rise to two long-range terms inversely proportional to l; one of these, like the Bjerknes force, is directed along the line joining the centers of the bubbles, and the other is in the direction of the gradient of an incident field. The refined expression of the interaction force is used for studies of the relative motion of the two bubbles, resulting from their radiative interaction, in a plane traveling wave and a plane standing wave. It is shown that the long-range terms can cause the bubbles to form stable bound pairs with a fixed separation between the bubbles. © *1997 Acoustical Society of America.* [S0001-4966(97)05108-4]

PACS numbers: 43.25.Yw [MAB]

INTRODUCTION

C. A. Bjerknes and his son V. F. K. Bjerknes were the first to take a serious interest in the radiative interaction of gas bubbles.[1–3] C. A. Bjerknes first derived an analytical expression for a time-averaged interaction force of two pulsating bubbles, which is since called in his honor the mutual or secondary Bjerknes force. It is important for the following that the above expression is based on a number of quite rigid assumptions, namely, that the surrounding liquid is incompressible and inviscid, the separation between the two bubbles is much larger than their radii, the bubbles keep a spherical shape during their pulsations, and the amplitude of the pulsations is small enough that nonlinear effects be negligible.

Subsequently, the mutual Bjerknes force has been the subject of a lot of experimental, theoretical, and numerical studies.[4–22] The works listed in Refs. 4–7, 9, and 11 essentially repeat Bjerknes' result. Kazantsev[8] and Crum[10] have carried out experiments that have shown satisfactory agreement with Bjerknes' theory. Zabolotskaya[12,13] attempted to investigate theoretically the interaction force for small separations between bubbles. This work was extended by the present authors[15,21] and yielded results that provided an explanation for the stable bubble clusters, ''bubble grapes,'' previously observed in experiments.[22] Oguz and Prosperetti[16] and Pelekasis and Tsamopoulos[17,18] conducted extensive numerical studies of the radiative interaction of two bubbles, allowing for nonlinear oscillations of the bubbles, and found significant departures from Bjerknes' theory for strong acoustic pressures. Among recent publications the works by Parlitz and collaborators[19,20] are of interest, where the interbubble forces are considered in the context of the bubble pattern formation in acoustic cavitation.

What all the above mentioned studies have in common is that as far as the interbubble forces are concerned, they start from the assumption of an incompressible liquid. Physi-cally, this means that separations between bubbles are considered to be much smaller than the wavelength of sound. Meanwhile, there are ultrasonic technologies and medical applications which use acoustic fields of very high frequencies (up to 10^9 Hz). In fields of this kind, separations between bubbles can be on the order of, or even more than, the wavelength of sound. As a result, the model of an incompressible liquid becomes illegitimate and the question arises: What happens to the radiative interaction force of two gas bubbles if the compressibility of the surrounding liquid is taken into account? As far as we are aware, the first attempt to answer this question was made by Nemtsov.[23] He calculated, taking into account the compressibility of the surrounding liquid, the interaction force of two gas bubbles in a plane traveling wave field and found that, first, in addition to the term obtained by C. A. Bjerknes (which is known to be inversely proportional to l^2, where l is the distance between the two bubbles), the interaction force involves a long-range term inversely proportional to l; second, the sign of the interaction force changes, depending on l; and third, the interaction force on one bubble is not equal and opposite to that on the other bubble (this result is accounted for by the fact that in a compressible liquid some part of the momentum is carried away by scattered waves to infinity). Thus it has been shown that the calculation of the interaction force from Bjerknes' formula in the cases where the compressibility of the surrounding liquid is no longer negligible can give both quantitatively and qualitatively wrong results. However, Nemtsov has lost an important term of the interaction force. The error has been revealed by the present authors[24] who have shown that the interaction force must involve one more long-range term directed along the wave vector of an incident field (the term obtained by Nemtsov, like the Bjerknes force, is directed along the line joining the centers of the bubbles). Also, dissipative losses have been taken into ac-

count and the case of a plane standing wave has been considered.[25]

In this paper, the goal is twofold. First, we seek to obtain the mutual interaction force of two gas bubbles in a compressible liquid for a sound field of arbitrary geometry and thereby to extend our previous results which are concerned with plane fields only. Second, we aim to study the relative motion of the two bubbles, resulting from their radiative interaction, and to reveal new phenomena due to the compressibility of the surrounding liquid. We will assume the surrounding liquid to be inviscid. The dissipative losses will be taken into account phenomenologically, by entering damping into equations of radial pulsations of the two bubbles, as is done in Ref. 12. In this connection, there is a point to be made. It is shown in an earlier work by one of the present authors[26] that the viscosity can vastly change the primary radiation force on an individual bubble. Most likely something of this kind happens to the secondary one as well. This, however, can occur provided the viscosity is strong enough, more specifically, when the viscous wavelength is large compared to the bubble radii. If this is not the case, then the above phenomenological approach is quite correct. We will assume here that the viscosity, as is usually the case in practice, is small enough for the phenomenological approach to be warranted.

I. EXPRESSION FOR THE INTERACTION FORCE IN A COMPRESSIBLE LIQUID

Consider two gas bubbles with the mean radii R_1 and R_2, suspended in a perfect compressible liquid and irradiated by a sound field. Suppose that R_1 and R_2 are much smaller than the wavelength λ of sound and the distance l between the equilibrium centers of the bubbles, whereas the ratio of l to λ is arbitrary. Thus there are four small parameters in the problem involved: kR_1, kR_2, R_1/l, and R_2/l, where k is the wave number. It is clear that the radiative interaction force can be developed as a series in powers of these parameters. It is the purpose of this section to find leading terms of this expansion.

The total sound field outside the bubbles can be represented as

$$\varphi = \varphi_I + \varphi_1 + \varphi_2, \tag{1}$$

where φ_I is the velocity potential of the incident field, and φ_α is the velocity potential of the field scattered by bubble number α ($\alpha = 1,2$).

The incident field can be written in the general case as

$$\varphi_I = a(\mathbf{r}) \exp(-i\omega t), \tag{2}$$

where $a(\mathbf{r})$ is the complex amplitude of the incident field at a space point given by the position vector $\mathbf{r}$. Note that $a(\mathbf{r})$ must satisfy the Helmholtz equation that Eq. (2) should actually describe a sound wave.

Assuming that the bubbles only execute radial pulsations, for the scattered field φ_α one has

$$\varphi_\alpha = a_\alpha h_0^{(1)}(kl_\alpha) \exp(-i\omega t), \tag{3}$$

where $h_0^{(1)}$ is the spherical Hankel function, $l_\alpha = |\mathbf{r} - \mathbf{r}_\alpha|$, $\mathbf{r}_\alpha$ is the position vector of the center of bubble α at rest, and

a_α is a constant to be determined by the boundary conditions at the surfaces of the bubbles. We must find a_α to zero order in kR_1 and kR_2 and up to first order in R_1/l and R_2/l. It will be shown below that this accuracy is sufficient to obtain leading terms of the interaction force.

The boundary condition for the liquid velocity $\mathbf{v}$ at the surface of bubble α is given by

$$\mathbf{v} \cdot \mathbf{n}_\alpha = \dot{x}_\alpha(t) \quad \text{at } l_\alpha = R_\alpha, \tag{4}$$

where $\mathbf{n}_\alpha$ is the outward normal to the surface of bubble α, $x_\alpha(t)$ is the change in the radius of bubble α at time t, and the dot over $x_\alpha(t)$ denotes differentiation with respect to t. Substituting Eqs. (2) and (3) into Eq. (1) and then Eq. (1) into Eq. (4), with the accuracy indicated above one obtains

$$a_\alpha \exp(-i\omega t) = -ikR_\alpha^2 \dot{x}_\alpha(t). \tag{5}$$

Here, we have also used the assumption that

$$|\nabla_\alpha a(\mathbf{r}_\alpha)| \ll \frac{|a_\alpha|}{kR_\alpha^2}, \tag{6}$$

where ∇_α denotes the gradient operator acting on the variable $\mathbf{r}_\alpha$. It will be shown below that this inequality is always satisfied provided $kR_{1,2} \ll 1$.

Substitution of Eq. (5) into Eq. (3) yields

$$\varphi_\alpha = -ikR_\alpha^2 \dot{x}_\alpha(t) h_0^{(1)}(kl_\alpha). \tag{7}$$

To find an explicit expression for $x_\alpha(t)$, one should apply the boundary conditions for the pressures outside and inside the bubbles. For bubble α one has

$$P_{g\alpha} = P_0 + P_I + P_1 + P_2 + P_{ST\alpha} \quad \text{at } l_\alpha = R_\alpha, \tag{8}$$

where $P_{g\alpha}$ is the pressure of the gas inside bubble α, given, assuming that the gas follows the adiabatic law, by

$$P_{g\alpha} = P_{g\alpha 0} \left(\frac{R_\alpha}{R_\alpha + x_\alpha(t)} \right)^{3\gamma} \approx P_{g\alpha 0} \left(1 - 3\gamma \frac{x_\alpha(t)}{R_\alpha} \right) \tag{9}$$

in which $P_{g\alpha 0}$ is the equilibrium gas pressure inside bubble α, and γ is the specific heat ratio; P_0 is the equilibrium pressure of the liquid; P_I is the change in the liquid pressure, produced by the incident wave and defined by

$$P_I = -\rho_0 \frac{\partial \varphi_I}{\partial t} \tag{10}$$

in which ρ_0 is the equilibrium density of the liquid; P_α ($\alpha = 1,2$) is the change in the liquid pressure, produced by the scattered wave from bubble α and given by

$$P_\alpha = -\rho_0 \frac{\partial \varphi_\alpha}{\partial t} \tag{11}$$

and, finally, $P_{ST\alpha}$ is the pressure of the surface tension given by the well-known formula

$$P_{ST\alpha} = \frac{2\sigma}{R_\alpha + x_\alpha(t)} \approx \frac{2\sigma}{R_\alpha} \left(1 - \frac{x_\alpha(t)}{R_\alpha} \right) \tag{12}$$

in which σ is the surface tension coefficient. Substituting Eq. (2) into Eq. (10), Eq. (7) into Eq. (11), and then Eqs. (9)–(12) into Eq. (8), one obtains for $P_{g\alpha 0}$

$$P_{g\alpha 0} = P_0 + \frac{2\sigma}{R_\alpha}, \tag{13}$$

and for $x_\alpha(t)$

$$\ddot{x}_\alpha + \omega_\alpha^2 x_\alpha + \omega\delta_\alpha \dot{x}_\alpha + \frac{R_\beta^2}{R_\alpha l}\exp(ikl)\ddot{x}_\beta$$

$$= -\frac{i\omega}{R_\alpha}a(\mathbf{r}_\alpha)\exp(-i\omega t), \tag{14}$$

where $\beta=2$ if $\alpha=1$ and vice versa. In order to take into account dissipation of energy, the term $\omega\delta_\alpha\dot{x}_\alpha$ has been inserted into Eq. (14), in which δ_α is the dimensionless damping due to heat, viscous, and radiation losses. An explicit expression for δ_α can be found, for example, in Ref. 27. It is obvious that solutions to Eq. (14) should be sought in the following form

$$x_\alpha(t) = C_\alpha \exp(-i\omega t). \tag{15}$$

Substituting Eq. (15) into Eq. (14) and performing calculations with the same accuracy as before, one finds

$$x_\alpha(t) = \frac{i \exp(-i\omega t)}{\omega R_\alpha(1-\omega_\alpha^2/\omega^2+i\delta_\alpha)}$$

$$\times \left(a(\mathbf{r}_\alpha) - \frac{R_\beta \exp(ikl)a(\mathbf{r}_\beta)}{l(1-\omega_\beta^2/\omega^2+i\delta_\beta)} \right). \tag{16}$$

We can now prove that inequality (6) is correct. By comparing Eq. (16) with Eq. (5), it can be easily seen that a_α is of the same order as the quantity $kR_\alpha a(\mathbf{r}_\alpha)$. Substitution of the latter into Eq. (6) leads us to conclude that the term on the right-hand side of Eq. (6) is of the order of $a(\mathbf{r}_\alpha)/R_\alpha$, while the term on the left-hand side of Eq. (6) is apparently of the order of $ka(\mathbf{r}_\alpha)$. Thus the ratio of the latter to the former is of the order of kR_α and hence inequality (6) does hold.

Now there is all that is needed to proceed to direct calculation of the interaction force. Let us make use of the well-known formula for the acoustic radiation force on a bubble (see, for example, Ref. 10) which in the case considered here takes the form

$$\mathbf{F}_\alpha = -\langle V_\alpha(t)\nabla(P_I(\mathbf{r},t)+P_\beta(\mathbf{r},t))|_{\mathbf{r}=\mathbf{r}_\alpha}\rangle, \tag{17}$$

where $V_\alpha(t)$ is the volume of bubble α at time t, given by

$$V_\alpha(t) = \tfrac{4}{3}\pi(R_\alpha + x_\alpha(t))^3 \tag{18}$$

or, with accuracy up to the first-order terms,

$$V_\alpha(t) \approx \frac{4}{3}\pi R_\alpha^3 \left(1+\frac{3x_\alpha(t)}{R_\alpha}\right), \tag{19}$$

and $\langle\ \rangle$ means an average over time. Note that the force $\mathbf{F}_\alpha$ contains the acoustic radiation pressure as well as the radiative interaction force. For this reason, it is appropriate to refer to $\mathbf{F}_\alpha$ as the total radiation force on a bubble (for the case at issue, on bubble α) in the presence of another bubble.

Using Eqs. (2), (7), (10), (11), and (19), after straightforward calculations, from Eq. (17) one obtains

$$\mathbf{F}_\alpha = \mathbf{F}_{P\alpha} + \mathbf{F}_{I\alpha}, \tag{20}$$

where

$$\mathbf{F}_{P\alpha} = -2\pi\rho_0 R_\alpha|s_\alpha|^{-2}\,\mathrm{Re}\{s_\alpha a^*(\mathbf{r}_\alpha)\nabla_\alpha a(\mathbf{r}_\alpha)\} \tag{21}$$

and

$$\mathbf{F}_{I\alpha} = \frac{2\pi\rho_0 R_1 R_2}{|s_1|^2|s_2|^2}\,\mathrm{Re}\left\{ s_\beta^* \exp(ikl)a(\mathbf{r}_\beta)\left(s_\alpha a^*(\mathbf{r}_\alpha)\frac{\mathbf{m}_{\alpha\beta}}{l^2}\right.\right.$$

$$\left.\left. -iks_\alpha a^*(\mathbf{r}_\alpha)\frac{\mathbf{m}_{\alpha\beta}}{l}+\frac{s_\alpha^*}{l}\nabla_\alpha a^*(\mathbf{r}_\alpha) \right)\right\} \tag{22}$$

in which

$$s_\alpha = 1-\omega_\alpha^2/\omega^2+i\delta_\alpha, \tag{23}$$

and $\mathbf{m}_{\alpha\beta} = (\mathbf{r}_\beta - \mathbf{r}_\alpha)/l$. It can be seen that the force $\mathbf{F}_{P\alpha}$ is the acoustic radiation force which would act on bubble α if the neighboring bubble were absent. It is usually called the acoustic radiation pressure or the primary Bjerknes force. The force $\mathbf{F}_{I\alpha}$ resulting from the mutual interaction of the bubbles with each other's scattered field is usually referred to as the radiative interaction force or the mutual interaction force of two bubbles in an acoustic field. It can be seen that in the limit involved the force $\mathbf{F}_{I\alpha}$ consists of three qualitatively different terms. The term inversely proportional to l^2 and directed along the line joining the centers of the bubbles is dominant in the limit $R_{1,2}\ll l\ll\lambda$, i.e., this is, ignoring the factor $\exp(ikl)$, the mutual Bjerknes force. Taking into account the compressibility of the surrounding liquid gives rise to two more terms, one of which is directed along the line of the centers of the bubbles, like the mutual Bjerknes force, and the other is directed along the gradient of the incident field. Both of these terms are inversely proportional to l and hence can prevail over the mutual Bjerknes force at large distances between the bubbles, at kl of the order of, or more than, unity to be exact. Besides, all three terms depend on the factor $\exp(ikl)$, and therefore the radiative interaction force can cause the bubbles to attract or repel each other depending on l.

Equations (20)–(23) can be applied to calculating the total radiation force in any sound field. However, plane traveling waves and plane standing waves are of prime interest in acoustic applications, and therefore it is these cases that will be considered below.

For a plane traveling wave, setting $a(\mathbf{r})=A\,\exp(i\mathbf{k}\cdot\mathbf{r})$, from Eqs. (20)–(22) one obtains

$$\mathbf{F}_\alpha^{(\mathrm{tr})} = \mathbf{F}_{P\alpha}^{(\mathrm{tr})} + \mathbf{F}_{I\alpha}^{(\mathrm{tr})}, \tag{24}$$

where

$$\mathbf{F}_{P\alpha}^{(\mathrm{tr})} = 2\pi\rho_0|A|^2 R_\alpha\delta_\alpha|s_\alpha|^{-2}\mathbf{k} \tag{25}$$

and

$$\mathbf{F}_{I\alpha}^{(\mathrm{tr})} = \frac{2\pi\rho_0|A|^2 R_1 R_2}{|s_1|^2|s_2|^2 l^2}\,[(B_1(\cos\chi_\alpha + kl\,\sin\chi_\alpha)$$

$$+(-1)^\alpha B_2(\sin\chi_\alpha - kl\,\cos\chi_\alpha))\mathbf{m}_{\alpha\beta}$$

$$+\mathbf{k}l(B_3\,\sin\chi_\alpha - B_4\,\cos\chi_\alpha)] \tag{26}$$

in which

$$B_1 = (1-\omega_1^2/\omega^2)(1-\omega_2^2/\omega^2)+\delta_1\delta_2, \tag{27}$$

$$B_2 = \delta_1(1 - \omega_2^2/\omega^2) - \delta_2(1 - \omega_1^2/\omega^2), \tag{28}$$

$$B_3 = (1 - \omega_1^2/\omega^2)(1 - \omega_2^2/\omega^2) - \delta_1\delta_2, \tag{29}$$

$$B_4 = \delta_1(1 - \omega_2^2/\omega^2) + \delta_2(1 - \omega_1^2/\omega^2), \tag{30}$$

$$\chi_\alpha = kl + \mathbf{k}\cdot\mathbf{m}_{\alpha\beta}l. \tag{31}$$

For a plane standing wave, setting $a(\mathbf{r}) = A\cos(\mathbf{k}\cdot\mathbf{r})$, one finds

$$\mathbf{F}_\alpha^{(\mathrm{st})} = \mathbf{F}_{P\alpha}^{(\mathrm{st})} + \mathbf{F}_{I\alpha}^{(\mathrm{st})}, \tag{32}$$

where

$$\mathbf{F}_{P\alpha}^{(\mathrm{st})} = \pi\rho_0|A|^2 R_\alpha|s_\alpha|^{-2}(1 - \omega_\alpha^2/\omega^2)\sin(2\mathbf{k}\cdot\mathbf{r}_\alpha)\mathbf{k} \tag{33}$$

and

$$\mathbf{F}_{I\alpha}^{(\mathrm{st})} = \frac{2\pi\rho_0|A|^2 R_1 R_2}{|s_1|^2|s_2|^2 l^2}\big[\mathbf{m}_{\alpha\beta}\cos(\mathbf{k}\cdot\mathbf{r}_1)\cos(\mathbf{k}\cdot\mathbf{r}_2)$$
$$\times(B_1(\cos kl + kl\sin kl) + (-1)^\alpha B_2(\sin kl$$
$$- kl\cos kl)) - \mathbf{k}\cos(\mathbf{k}\cdot\mathbf{r}_\beta)\sin(\mathbf{k}\cdot\mathbf{r}_\alpha)$$
$$\times(B_3\cos kl + B_4\sin kl)\big]. \tag{34}$$

Next, the expressions for the total radiation force derived in this section are used for studies of the relative motion of two gas bubbles in high-frequency acoustic fields where λ is comparable to l.

II. RELATIVE MOTION OF TWO GAS BUBBLES

When placed in an acoustic field, gas bubbles oscillate with the driving frequency. These oscillations may be conventionally referred to as ''fast.'' While calculating the acoustic radiation forces in the preceding section, it is over the ''fast'' oscillations that we averaged equations. At the same time, the bubbles execute ''slow'' translational motions due to the radiation forces and also some others, such as buoyancy, drag, etc. It is the purpose of this section to study these ''slow'' motions. To put it more exactly, we shall be interested in new phenomena that appear in the relative motion of two gas bubbles as a result of the long-range terms of the interaction force.

The ''slow'' translational motion of bubble α in an acoustic field is given by

$$(m_{g\alpha} + M_\alpha)\frac{d\mathbf{U}_\alpha}{dt} = \mathbf{F}_{G\alpha} + \mathbf{F}_{A\alpha} + \mathbf{F}_{D\alpha} + \mathbf{F}_\alpha, \tag{35}$$

where $m_{g\alpha}$ is the mass of the gas inside bubble α, $M_\alpha = 2\pi\rho_0 R_\alpha^3/3$ is the so-called added or apparent mass, $\mathbf{U}_\alpha$ is the ''slow'' translational velocity of bubble α, $\mathbf{F}_{G\alpha} = m_{g\alpha}\mathbf{g}$ is the gravity, $\mathbf{F}_{A\alpha} = -4\pi\rho_0 R_\alpha^3\mathbf{g}/3$ is the buoyant force, and $\mathbf{F}_{D\alpha}$ is the viscous drag force which is assumed to follow Stokes' law

$$\mathbf{F}_{D\alpha} = -6\pi\eta R_\alpha\mathbf{U}_\alpha, \tag{36}$$

where η is the viscosity of the liquid. The inertial term, i.e., the term on the left-hand side of Eq. (35), can be dropped since in cases of interest it is, as a rule, much less than the others.[10] As a result, Eq. (35) becomes

$$\mathbf{U}_\alpha = \frac{d\mathbf{r}_\alpha}{dt} = \frac{\mathbf{F}_{G\alpha} + \mathbf{F}_{A\alpha} + \mathbf{F}_\alpha}{6\pi\eta R_\alpha}. \tag{37}$$

By subtracting Eq. (37) at $\alpha = 1$ from Eq. (37) at $\alpha = 2$, one obtains an equation of the relative motion of the bubbles to be

$$\frac{d\mathbf{l}}{dt} = \frac{1}{6\pi\eta}\left(\frac{\mathbf{F}_2}{R_2} - \frac{\mathbf{F}_1}{R_1}\right) + \frac{2\mathbf{g}}{9\eta}((\rho_2 - \rho_0)R_2^2 - (\rho_1 - \rho_0)R_1^2), \tag{38}$$

where $\mathbf{l} = \mathbf{r}_2 - \mathbf{r}_1$, and ρ_1 and ρ_2 are the equilibrium densities of the gas inside bubbles 1 and 2, respectively.

A. Plane traveling wave

Substituting Eqs. (24)–(26) into Eq. (38) and putting the equation obtained into a nondimensional form, one finds

$$\frac{d\mathbf{q}}{d\tau} = \mathbf{S} - \frac{\mathbf{e}_k}{B_1\xi}(\sin x(B_3\cos\xi + B_4\sin\xi)$$
$$- \mu\cos x(B_3\sin\xi - B_4\cos\xi)) - \frac{\mathbf{e}_l}{B_1\xi^2}((\cos\xi$$
$$+ \xi\sin\xi)(B_1\cos x - B_2\sin x) + \mu(\sin\xi$$
$$- \xi\cos\xi)(B_1\sin x + B_2\cos x)), \tag{39}$$

where $\mathbf{q} = k\mathbf{l}$, $\xi = |\mathbf{q}| = kl$ is the dimensionless distance between the bubbles, $\mathbf{e}_k = \mathbf{k}/k$, $\mathbf{e}_l = \mathbf{l}/l$, $x = \xi(\mathbf{e}_k\cdot\mathbf{e}_l)$ is the projection of $\mathbf{q}$ on $\mathbf{k}$, $\mu = (R_1 - R_2)/(R_1 + R_2)$, $\tau = Dt$ is the dimensionless time, and D is the positive constant given by

$$D = \frac{\rho_0|A|^2 k^3(R_1 + R_2)B_1}{3\eta|s_1|^2|s_2|^2}, \tag{40}$$

$\mathbf{S}$ is the constant vector defined by

$$\mathbf{S} = S_k\mathbf{e}_k + S_g\mathbf{e}_g, \tag{41}$$

where

$$S_k = \frac{\delta_2|s_1|^2 - \delta_1|s_2|^2}{k(R_1 + R_2)B_1}, \tag{42}$$

$$S_g = \frac{2g|s_1|^2|s_2|^2((\rho_2 - \rho_0)R_2^2 - (\rho_1 - \rho_0)R_1^2)}{3\rho_0|A|^2 k^2(R_1 + R_2)B_1}, \tag{43}$$

and $g = |\mathbf{g}|$, $\mathbf{e}_g = \mathbf{g}/g$.

The intricate structure of the right-hand side of Eq. (39) makes it impossible to find an analytical solution to this equation in the general case. However, some quite interesting information can be still extracted. The analysis of Eq. (39) suggests that under certain conditions the bubbles can unite to form a stable bound pair in which the separation between the bubbles is kept constant as long as the bubbles are in a sound field. Thus the pair moves as a single entity. In order to ascertain conditions under which this is the case, one should put $d\mathbf{q}/d\tau = 0$ in Eq. (39) following which one obtains

$$\frac{\mathbf{e}_l}{B_1}((\cos\,\xi+\xi\,\sin\,\xi)(B_1\,\cos\,x-B_2\,\sin\,x)$$

$$+\mu(\sin\,\xi-\xi\,\cos\,\xi)(B_1\,\sin\,x+B_2\,\cos\,x))$$

$$+\frac{\mathbf{e}_k\xi}{B_1}\,(\sin\,x(B_3\,\cos\,\xi+B_4\,\sin\,\xi)-\mu\,\cos\,x(B_3\,\sin\,\xi$$

$$-B_4\,\cos\,\xi))=\xi^2\mathbf{S}. \tag{44}$$

From Eq. (44) it follows that while the bubbles form the bound pair, the vector $\mathbf{e}_l$ is in the same plane as the vectors $\mathbf{e}_k$ and $\mathbf{e}_g$. Besides, if the bubbles are of appreciably different radii, then the magnitude of $\mathbf{S}$ is large compared to unity in cases of interest. For example, for air bubbles in water, if we take $R_1=20\ \mu\mathrm{m}$, $R_2=40\ \mu\mathrm{m}$, $f=130\ \mathrm{kHz}$, and $\Delta P_I=1\ \mathrm{kPa}$ (where $f=2\pi/\omega$ and $\Delta P_I=\rho_0\omega|A|$ are the frequency and the pressure amplitude of the incident wave, respectively), we obtain: $\mathbf{S}\approx8\mathbf{e}_k+72\mathbf{e}_g$. Taking this fact into consideration and comparing the left-hand and right-hand sides of Eq. (44) with each other, we conclude that bubbles of appreciably different radii can form the bound pair only when $\xi\ll1$. Note that in this limit the long-range terms become negligible and hence the radiative interaction force turns into the mutual Bjerknes force. Thus the bound states of this kind are formed due to counteraction between the buoyancy, the radiation pressure, and the mutual Bjerknes force.

Passing to the limit $\xi\ll1$ and retaining only the leading terms, from Eq. (44) one finds

$$\mathbf{e}_l=\xi^2\mathbf{S}. \tag{45}$$

Thus the relative position of the bubbles when they are in the bound state is defined by

$$\xi=|\mathbf{S}|^{-1/2},\quad x=(\mathbf{S}\cdot\mathbf{e}_k)|\mathbf{S}|^{-3/2}. \tag{46}$$

To investigate the stability of this state for disturbances, the well-known method suggested by Lyapunov may be used. For this purpose, one should return to Eq. (39) and pass to the limit $\xi\ll1$. As a result, one obtains

$$\frac{d\mathbf{q}}{d\tau}=\mathbf{S}-\frac{\mathbf{e}_l}{\xi^2}. \tag{47}$$

It is convenient to introduce a coordinate system with the x axis parallel to the vector $\mathbf{e}_k$, and the y axis normal to it and lying in the same plane as the vector $\mathbf{e}_l$ (recall that when the bubbles are in the bound state the vectors $\mathbf{e}_k$, $\mathbf{e}_l$, and $\mathbf{e}_g$ lie in the same plane). Equation (47) can then be represented as a set of two equations

$$\frac{dx}{d\tau}=S_x-\frac{x}{\xi^3},\quad\frac{dy}{d\tau}=S_y-\frac{y}{\xi^3}, \tag{48}$$

where y is the projection of ξ on the y axis, i.e., $\xi=(x^2+y^2)^{1/2}$, and S_x and S_y are the projections of $\mathbf{S}$ on the x and y axes, respectively. These equations can be put into more convenient form

$$\frac{dx}{d\tau}=\frac{\partial F(x,y)}{\partial x},\quad\frac{dy}{d\tau}=\frac{\partial F(x,y)}{\partial y}, \tag{49}$$

where $F(x,y)=S_x x+S_y y+1/\xi$. Assume now that the bubbles were initially in the bound state given by Eqs. (46) and then were disturbed from the equilibrium, their current position being determined by $x=\bar{x}+x'(\tau)$ and $y=\bar{y}+y'(\tau)$, where $\bar{x}$ and $\bar{y}$ are the values of x and y in the bound state, and x' and y' are small deviations. Substituting these values of x and y into Eqs. (49) and retaining only the first-order terms in x' and y', one obtains

$$\frac{dx'}{d\tau}=a_{11}x'+a_{12}y',$$
$$\frac{dy'}{d\tau}=a_{21}x'+a_{22}y', \tag{50}$$

where the coefficients a_{ij} are defined by

$$a_{11}=\frac{\partial^2 F(\bar{x},\bar{y})}{\partial\bar{x}^2},\quad a_{12}=a_{21}=\frac{\partial^2 F(\bar{x},\bar{y})}{\partial\bar{x}\partial\bar{y}},$$

$$a_{22}=\frac{\partial^2 F(\bar{x},\bar{y})}{\partial\bar{y}^2}. \tag{51}$$

The set of Eqs. (50) is equivalent to

$$\frac{d^2x'}{d\tau^2}-(a_{11}+a_{22})\,\frac{dx'}{d\tau}+(a_{11}a_{22}-a_{12}^2)x'=0. \tag{52}$$

It is clear that a solution to Eq. (52) should be sought in the form $x'=C\,\exp(\nu\tau)$. Substitution of this expression into Eq. (52) yields

$$\nu^2-(a_{11}+a_{22})\nu+a_{11}a_{22}-a_{12}^2=0, \tag{53}$$

from which one finds

$$\nu_{1,2}=\tfrac{1}{2}(a_{11}+a_{22})\pm(\tfrac{1}{4}(a_{11}+a_{22})^2-a_{11}a_{22}+a_{12}^2)^{1/2}. \tag{54}$$

For the bound state of the bubbles to be stable, the real parts of both roots of Eq. (53) must be negative. Substitution of Eqs. (46) into Eqs. (51) and then into Eq. (54) results in $\nu_1=2\xi^{-3}$ and $\nu_2=-\xi^{-3}$. This fact leads us to conclude that bubbles of appreciably different radii cannot form a stable bound pair at any rate if the separation between them is large compared with their sizes as supposed in this study.

Consider now bubbles of close radii. The complexity of Eq. (44) does not enable one to find analytical solutions without simplifying assumptions. Therefore to show that for these bubbles stable bound states do exist, consider the following case. Assume that, first, $R_1=R_2$, and second, the resonance frequency of the bubbles is appreciably different from the frequency of the incident field. The second assumption allows us to neglect damping and put $\delta_1=\delta_2=0$ in Eq. (44). As a result, Eq. (44) becomes

$$\mathbf{e}_l\,\cos\,x(\cos\,\xi+\xi\,\sin\,\xi)+\mathbf{e}_k\xi\,\sin\,x\,\cos\,\xi=0. \tag{55}$$

By introducing a coordinate system similar to that used above, instead of Eq. (55) one obtains

$$x\,\cos\,x(\cos\,\xi+\xi\,\sin\,\xi)+\xi^2\,\sin\,x\,\cos\,\xi=0,$$
$$y\,\cos\,x(\cos\,\xi+\xi\,\sin\,\xi)=0. \tag{56}$$

It can be easily seen that Eqs. (56) are equivalent to three sets of simpler equations

$$\cos x = 0, \quad \cos \xi = 0, \tag{57}$$

$$\sin x = 0, \quad \cos \xi + \xi \sin \xi = 0, \tag{58}$$

and

$$y = 0, \quad \cos x + 2x \sin x = 0. \tag{59}$$

Equations (57) have exact solutions

$$\xi_n = (2n+1) \frac{\pi}{2}, \quad n = 0,1,2,...,$$
$$x_m = (2m+1) \frac{\pi}{2}, \quad -n \le m \le n. \tag{60}$$

The sets of Eqs. (58) and (59) can be solved approximately. From Eqs. (58) one finds

$$\xi_n \approx n\pi - \frac{1}{n\pi}, \quad n = 1,2,...,$$
$$x_m = m\pi, \quad -n < m < n. \tag{61}$$

Similarly, from Eqs. (59) one has

$$\xi_n = |x_n| \approx n\pi - \frac{1}{2n\pi}, \quad n = 1,2,... . \tag{62}$$

The larger the n in Eqs. (61) and (62), the more accurate are the values of ξ_n given by these equations.

To examine the stability of these states, apply Lyapunov's method again. By setting $R_1 = R_2$ and $\delta_1 = \delta_2 = 0$, from Eq. (39) one obtains

$$\frac{d\mathbf{q}}{d\tilde{\tau}} = -\mathbf{e}_k \frac{\sin x \cos \xi}{\xi} - \mathbf{e}_l \frac{\cos x(\cos \xi + \xi \sin \xi)}{\xi^2}, \tag{63}$$

where $\tilde{\tau} = \tilde{D}t$, and $\tilde{D}$ is the positive constant given by

$$\tilde{D} = \frac{2\rho_0 |A|^2 k^3 R_1}{3\eta(1 - \omega_1^2/\omega^2)}. \tag{64}$$

By writing Eq. (63) in terms of the projections on the axes of the coordinate system introduced above, one finds

$$\frac{dx}{d\tilde{\tau}} = -\frac{\sin x \cos \xi}{\xi} - \frac{x \cos x(\cos \xi + \xi \sin \xi)}{\xi^3}$$
$$\frac{dy}{d\tilde{\tau}} = -\frac{y \cos x(\cos \xi + \xi \sin \xi)}{\xi^3}, \tag{65}$$

or, putting these equations into a more compact and convenient form,

$$\frac{dx}{d\tilde{\tau}} = \frac{\partial \Phi(x,y)}{\partial x}, \quad \frac{dy}{d\tilde{\tau}} = \frac{\partial \Phi(x,y)}{\partial y}, \tag{66}$$

where $\Phi(x,y) = \xi^{-1} \cos x \cos \xi$. The further operations are identical to those performed in deriving Eqs. (50)–(54). A difference is that the function $F(x,y)$ in Eqs. (51) should be replaced by $\Phi(x,y)$. Taking this into consideration and substituting Eqs. (60) into Eq. (54), one obtains

$$\nu_{1,2} = \xi_n^{-2}((-1)^{n+m} x_m \pm \xi_n)). \tag{67}$$

It can be easily seen that the roots are not negative simultaneously and hence the bound states defined by Eqs. (60) are unstable.

In a similar manner, for Eqs. (61) one finds

$$\nu_{1,2} \approx -\xi_n^{-2}((-1)^{n+m}\xi_n \pm |x_m|). \tag{68}$$

It follows that both roots are negative only when the sum $n+m$ is even. Accordingly, the corresponding bound states are stable.

Finally, for Eq. (62) one has

$$\nu_1 \approx -\frac{4\xi_n^2 - 1}{2\xi_n^3}, \quad \nu_2 \approx -\frac{1}{2\xi_n^3}. \tag{69}$$

Both roots are negative. Hence these bound states are stable.

To summarize, it should be expected that in reality the bound states can be observed given either by Eqs. (61) with the even sum $n+m$ or by Eq. (62). Note also that the stable bound state with the minimum separation between the bubbles is given by Eq. (62) at $n=1$, the separation being approximately equal to $\lambda/2$. In view of this, one can conclude that the compressibility of the liquid can manifest itself even at distances between the bubbles that are perceptibly smaller than the wavelength of an incident field.

B. Plane standing wave

One can see from Eqs. (33) and (34) that the ratio of the force $\mathbf{F}_{I\alpha}^{(st)}$ to the force $\mathbf{F}_{P\alpha}^{(st)}$ is of the order of (kR_β) $((kl)^2(1 - \omega_\beta^2/\omega^2))^{-1}$ if the pair of the bubbles at issue is sufficiently far from planes for which $\mathbf{k} \cdot \mathbf{r} = n\pi/2$, $n = 0$, $\pm 1, \pm 2,...$ (the force $\mathbf{F}_{P\alpha}^{(st)}$ becomes vanishingly small near these planes). As supposed l is of the order of the wavelength. It follows that the force $\mathbf{F}_{P\alpha}^{(st)}$ far exceeds the force $\mathbf{F}_{I\alpha}^{(st)}$ under the above condition. This conclusion is corroborated by experiments.[13] Assume for simplicity that $\mathbf{k} \perp \mathbf{g}$ (this case is often met with in practice). Then, if there is an aggregate of gas bubbles, it should be expected that the bubbles will first gather at the planes in which the radiation pressure (i.e., the force $\mathbf{F}_{P\alpha}^{(st)}$) vanishes. From Eq. (33) it follows that large bubbles, i.e., ones with the resonance frequencies less than the driving frequency, must accumulate at the planes for which n is odd. These planes correspond to the nodes of the acoustic pressure, and therefore, after getting to them, the large bubbles practically cease pulsating. As a result, both the radiation pressure and the interaction forces of the bubbles vanish. Thus the translational motions of the bubbles are defined mainly by the buoyancy and Stokes' drag force. We are not interested in motions of this kind in the present paper.

Small bubbles, i.e., those for which $\omega_\alpha > \omega$, must accumulate at the planes with even values of n. These planes pass through the antinodes of the acoustic wave pressure, and therefore the pulsations of the small bubbles do not vanish, conversely, they become maximal.

Suppose that the above "sorting" of bubbles has already occurred and consider the relative motion of two small bubbles. Substituting Eqs. (32)–(34) into Eq. (38), setting $\mathbf{k} \cdot \mathbf{r}_1 = \mathbf{k} \cdot \mathbf{r}_2 = n\pi$ and putting the equation obtained into a nondimensional form, one has

$$\frac{d\mathbf{q}}{d\tau} = S_g \mathbf{e}_g - \mathbf{e}_l \frac{\cos \xi + \xi \sin \xi + \zeta(\sin \xi - \xi \cos \xi)}{\xi^2},$$

(70)

where $\zeta = \mu B_2 / B_1$. As before, we are interested in bound states of the bubbles. From Eq. (70) it follows that the values of ξ at which the bubbles form a bound pair must satisfy the following equation

$$\mathbf{e}_l(\cos \xi + \xi \sin \xi + \zeta(\sin \xi - \xi \cos \xi)) = \xi^2 S_g \mathbf{e}_g. \quad (71)$$

As already mentioned, for bubbles of appreciably different radii the magnitude of S_g is large compared to unity. Therefore the bubbles of this kind can form a bound pair only if $\xi \ll 1$. For this case, from Eq. (71) one finds

$$\mathbf{e}_l = \frac{S_g}{|S_g|} \mathbf{e}_g \quad \text{and} \quad \xi = |S_g|^{-1/2}. \quad (72)$$

It can be readily verified, following Lyapunov's method, that, as is the case with a traveling wave, this bound state is unstable.

Consider now bubbles of close radii. As before, to simplify the problem, we assume the radii of the bubbles to be exactly equal to each other. As a result, Eq. (71) yields

$$\cos \xi + \xi \sin \xi = 0, \quad (73)$$

approximate solutions to which are

$$\xi_n \approx n\pi - \frac{1}{n\pi}, \quad n = 1,2,\dots. \quad (74)$$

These bound states are stable at even values of n. Thus the stable state with the minimum separation between the bubbles must be observed at $l \approx \lambda$. As is the case with a traveling wave, the stable bound states of bubbles in a standing wave field are due to the long-range terms of the interaction force.

III. CONCLUSIONS

This paper sought to answer the following question: What happens to the radiative interaction force of two gas bubbles when the wavelength of an incident acoustic field is comparable to the distance between the bubbles? In this case the compressibility of the surrounding liquid is no longer negligible and as a consequence the applicability of Bjerknes' theory, based on the assumption of an incompressible liquid, becomes questionable. It has been shown that when it is taken into account, the compressibility of the liquid gives rise to two long-range terms which, unlike the mutual Bjerknes force, are inversely proportional to the distance between the bubbles, not the distance squared; one of these terms, like the mutual Bjerknes force, is directed along the center line of the bubbles, while the other is in the direction of the gradient of an incident field.

The newly obtained expression for the interaction force has been applied to studying the "slow" relative motion of two gas bubbles in a plane traveling wave and a plane stand-

ing wave. It has been found that the long-range terms substantially affect this motion, in particular, due to those the bubbles can form stable bound pairs with a fixed separation between the bubbles. The results obtained may be useful in applications dealing with the radiative interaction of gas bubbles in high-frequency acoustic fields. It is of interest to note that long-range radiative forces similar to those considered here also exist in electrodynamics: they arise between two charged particles or two magnetic moments irradiated by an electromagnetic wave.[28,29]

[1] C. A. Bjerknes, *Hydrodynamische Fernkrafte* (Engelmann, Leipzig, 1915).

[2] V. F. K. Bjerknes, *Fields of Force* (Columbia U. P., New York, 1906).

[3] V. F. K. Bjerknes, *Die Kraftfelder* (Vieweg und Sohn, Braunschweig, 1909).

[4] W. Hicks, "On the problem of two pulsating spheres in a fluid," Proc. Camb. Philos. Soc. **3**, 276–285 (1879).

[5] W. Hicks, "On the problem of two pulsating spheres in a fluid (part II)," Proc. Camb. Philos. Soc. **4**, 29–35 (1880).

[6] K. Pearson, "On the motion of spherical and ellipsoidal bodies in fluid media," Q. Appl. Math. **20**, 60–80 (1884).

[7] A. B. Basset, *A Treatise on Hydrodynamics* (Cambridge U. P., London, 1888), p. 229.

[8] V. F. Kazantsev, "The motion of gas bubbles in a liquid under action of Bjerknes forces arising in an acoustic field," Dokl. Akad. Nauk SSSR **129**, 64–67 (1959).

[9] G. N. Kuznetsov and I. E. Shchekin, "Interaction of pulsating bubbles in a viscous fluid," Akust. Zh. **18**, 565–570 (1972).

[10] L. A. Crum, "Bjerknes forces on bubbles in a stationary sound field," J. Acoust. Soc. Am. **57**, 1363–1370 (1975).

[11] V. N. Alekseev, "Force produced by the acoustic radiation pressure on a sphere," Sov. Phys. Acoust. **29**, 77–81 (1983).

[12] E. A. Zabolotskaya, "Interaction of gas bubbles in a sound field," Sov. Phys. Acoust. **30**, 365–368 (1984).

[13] E. A. Zabolotskaya, "Nonlinear waves in a fluid with gaseous bubbles," Tr. Inst. Ob. Fiz. Akad. Nauk **18**, 121–155 (1989).

[14] *Physical Principles of Medical Ultrasonics*, edited by C. R. Hill (Ellis Horwood, New York, 1986).

[15] S. T. Zavtrak, "On the Bjerknes interaction force of two gas bubbles in a sound wave field," Akust. Zh. **33**, 240–245 (1987).

[16] H. N. Oguz and A. Prosperetti, "A generalization of the impulse and virial theorems with an application to bubble oscillations," J. Fluid Mech. **218**, 143–162 (1990).

[17] N. A. Pelekasis and J. A. Tsamopoulos, "Bjerknes forces between two bubbles. Part 1. Response to a step change in pressure," J. Fluid Mech. **254**, 467–499 (1993).

[18] N. A. Pelekasis and J. A. Tsamopoulos, "Bjerknes forces between two bubbles. Part 2. Response to an oscillatory pressure field," J. Fluid Mech. **254**, 501–527 (1993).

[19] I. Akhatov, U. Parlitz, and W. Lauterborn, "Pattern formation in acoustic cavitation," J. Acoust. Soc. Am. **96**, 3627–3635 (1994).

[20] U. Parlitz, C. Scheffczyk, I. Akhatov, and W. Lauterborn, "Structure formation in cavitation bubble fields," Chaos Solitons Fractals **5**, 1881–1891 (1995).

[21] A. A. Doinikov and S. T. Zavtrak, "On the mutual interaction of two gas bubbles in a sound field," Phys. Fluids **7**, 1923–1930 (1995).

[22] A. A. Doinikov and S. T. Zavtrak, "On the bubble grapes induced by a sound field," J. Acoust. Soc. Am. **99**, 3849–3850 (1996).

[23] B. E. Nemtsov, "Effects of the radiation interaction of bubbles in a fluid," Pis'ma Zh. Tekh. Fiz. **9**, 858–861 (1983).

[24] A. A. Doinikov and S. T. Zavtrak, "Radiation effects of the interaction of two gas bubbles in a compressible nonideal fluid," Akust. Zh. **36**, 429–432 (1990).

[25] A. A. Doinikov and S. T. Zavtrak, "Radiation interaction of gas bubbles

in a compressible fluid in a nonhomogeneous sound wave field,'' Pis'ma Zh. Tekh. Fiz. **16**, 63–66 (1990).

[26] A. A. Doinikov, ''Acoustic radiation pressure on a compressible sphere in a viscous fluid,'' J. Fluid Mech. **267**, 1–21 (1994).

[27] C. S. Clay and H. Medwin, *Acoustical Oceanography* (Wiley, New York, 1977).

[28] S. T. Zavtrak, ''A classical treatment of the long-range radiative interaction of small particles,'' J. Phys. A: Math. Gen. **23**, 1493–1499 (1990).

[29] S. T. Zavtrak, ''Long-range radiative interaction of magnetic moments in an electromagnetic wave,'' J. Phys. A: Math. Gen. **23**, L599–L602 (1990).

Models and measurements of sound propagation from a point source over mixed impedance ground

Patrice Boulanger, Tim Waters-Fuller, Keith Attenborough, and Kai Ming Li
Engineering Mechanics, Faculty of Technology, The Open University, Milton Keynes MK7 6AA, United Kingdom

Measurements of the excess attenuation of sound from a point source over mixed impedance ground in an anechoic chamber are compared with predictions obtained from models based on (a) the semiempirical theory due to De Jong, (b) Nyberg's theory, (c) a Fresnel-zone approximation, and (d) a boundary element code. The impedance discontinuities studied in this work are perpendicular to the source–receiver line. When there is a single discontinuity between acoustically hard and finite impedance surfaces, the De Jong semiempirical model, the Fresnel-zone model and boundary element code are found to give satisfactory agreement with measured data. The frequency of the first maximum in attenuation is found to be highest at approximately 70% hard surface cover rather than at the 100% expected. It is argued that this is a result of edge diffraction. When extended to multiple impedance discontinuities, the De Jong semiempirical model performs poorly. However, both Nyberg's theory and the boundary element code give good agreement with measured results, and the Fresnel-zone model gives qualitative agreement with the measured data. The measurements confirm that, under certain constraints given by Nyberg's theory, the ground effect due to mixed impedance may be determined from that predicted by using the area-average impedance. © *1997 Acoustical Society of America.* [S0001-4966(97)05409-X]

PACS numbers: 43.28.Fp, 43.58.Bh [LCS]

INTRODUCTION

The propagation of sound above ground of mixed impedance is a subject of increasing interest for noise prediction and for exploitation of ground effect. A number of studies have been presented in the past in order to develop models for predicting the effect of mixed impedance on sound propagation over flat ground.

Naghieh and Hayek[1] have developed an analytical solution to the problem of noise propagation from a point source above two infinite half-planes of different impedance. Their solution is based on wedge diffraction theory where the diffracted field of an impedance transition is calculated by taking it as a wedge with a top angle of 180° and different impedance on each side. The theory predicts the existence of two surface waves for the two half-planes, however, their solution is not straightforward and requires numerical computation. Enflo and Enflo[2] gave an exact solution for the singular integral equation obtained from the Helmholtz equation and the boundary condition when the source and receiver are close to the ground and many wavelengths away from the impedance discontinuity. Naze Tjøtta[3] found a simple analytical solution for a pure tone, a transient wave, and for the impulse response and showed numerical results illustrating the structure of the sound field both in the near field and the far field.

Apart from the theories presented above, the models developed to date can be classified into two categories. The first category is based on numerical methods. Durnin and Bertoni[4] used two approaches based on Green's theorem and obtained excellent agreement between their predictions and measurements of the excess attenuation at 1 kHz over ground presenting various proportions of grass and concrete in the case of source and receiver located at ground level. Harriott et al.[5] developed an approximate method for solving the boundary integral equations for propagation from a point source at a single frequency and predicted simple trends for a broadband A-weighted typical environmental noise source. A model developed by Rasmussen[6] involves joining the two parts of the sound field calculated, respectively, over the two infinite half-planes at the impedance discontinuity. The predictions of excess attenuation for a loudspeaker were in good agreement with the experiments. A wide-angle parabolic method (PE) in two dimensions was developed by Galindo[7] for sound propagation through a homogeneous atmosphere over mixed impedance ground. The predictions of the model were compared with a model based on ray theory and with measurements made in a scale model. The results of the predictions showed that the PE method is a powerful method of predicting sound-pressure level at frequencies between 100 and 2.5 kHz when sound propagates over a porous strip in a hard surface. The NASA implicit finite-difference (NIFD) implementation of the parabolic equation was used by Robertson et al.[8] to predict long-range, low-frequency sound propagation over varying terrain. A pair of test problems demonstrated the accuracy of the implementation, and predictions of pure-tone signals were made for propagation paths including portions of a flat lake surface in both quiescent and windy atmospheres.

A second category of models is based on semiempirical solutions. Koers[9] presented a model based on Kirchhoff diffraction theory where the diffracted field at an impedance transition is calculated by taking it as a wedge with a top angle of 180° and different impedance on each side. The results shown for two different proportions between hard and

soft surfaces compare well with measurements. Daigle *et al.*[10] compared indoor and outdoor measurements of excess attenuation to semiempirical solutions and solutions from Green's theorem for various proportions of hard and soft surfaces for a given geometry and showed that the semiempirical models fail when source and receiver are near the ground. More recently Hothersall and Harriot[11] have compared predictions of the approximate numerical method based on solving the boundary integral equations to those of both the semiempirical method developed by De Jong[12] and a third model where the boundary is simply related to the region around the specular reflection point defined by a Fresnel-zone condition. The result of their study performed for single or double discontinuity cases are presented for a few frequencies as graphs of excess attenuation versus the distance between source and impedance discontinuity. The agreement between the De Jong model and the boundary element method was generally good except near grazing. The Fresnel-zone method was found to be adequate in describing the general trends in the results of the other methods.

This review of previous work on sound propagation above mixed impedance ground shows clearly that there is a need for more validation of model predictions of excess attenuation spectra in the case of both single and multiple impedance discontinuities. The latter situation has received relatively little attention in the literature. Another point of interest is the investigation of the influence that various proportions of mixed impedance might have on ground effect (through variation in the frequency and phase of the first interference dip). Moreover, it is interesting to compare the very different cases of an impedance plane containing a single discontinuity and a striped impedance plane. The objective of this paper is to contribute some of the necessary additional work.

Four specific models are investigated and compared to measured data.

(1) The Fresnel-zone model

The Fresnel-zone model is the simplest and can be applied to a single impedance discontinuity or to multiple impedance discontinuities. It assumes that the reflecting area in a discontinuous plane is related simply to the region around the specular reflection point defined by a Fresnel-zone condition. Following Hothersall and Harriot[11] we assume that the excess attenuation is dependent on the proportion of the different surface areas inside the elliptical Fresnel zone and we improve the method initially by deriving the correct boundary values of the ellipse. A second modification involves the introduction of a new definition of the excess attenuation in mixed impedance cases. Instead of a linear interpolation between the predicted excess attenuation above uniform boundaries of each impedance, we use a linear interpolation between the pressure at the receiver relative to the free field above a surface of impedance Z_1 and the pressure at the receiver relative to the free field above a surface of impedance Z_2. Predictions obtained by this modified model show a better agreement with our measured data than the results calculated from the previous model.

(2) The De Jong model

The De Jong model[12] consists of using an empirical modification of the expression for diffraction at a rigid half-plane that can be easily applied to a single discontinuity and extended to multiple discontinuities.

(3) Nyberg's theory

More recently, Nyberg[13] presented a solution of the Helmholtz equation for multiple discontinuities using a Fourier transform technique. He showed that under certain conditions the ground effect due to mixed impedance may be determined from that predicted by using the area-averaged impedance.

(4) The boundary element technique

The fourth method was developed by Chandler-Wilde and Hothersall[14] and relies on a numerical solution of the problem of propagation from a line source by use of a boundary element technique. It has the advantage of allowing for any number of discontinuities.

The impedance model incorporated into these four approaches assumes a microstructure of a log-normal distribution of triangular pores sizes.[15] The model has been simplified for the repetitive calculations necessary here by the use of Padé approximation techniques.[16,17]

I. THEORIES

A. Theories of sound propagation over a single impedance discontinuity

1. The De Jong model

This model has been developed by De Jong[12] and uses an approximation of diffraction theory at a semi-infinite wedge. Taking the limit of zero angle for the wedge and obtaining a semi-infinite screen, De Jong presented a heuristic solution by analogy between the case of a semi-infinite screen, which can be regarded as an admittance step from an acoustically hard surface to air, with the case of an admittance step occurring at a transition between two types of ground.

In all further calculations the excess attenuation is obtained from

$$EA = 20 \log \left| \frac{P_1}{P} \right|, \tag{1}$$

where

$$P = P_0 \left[\frac{\exp(ikR_1)}{R_1} + Q \frac{\exp(ikR_2)}{R_2} \right] \tag{2}$$

is the total pressure at the receiver due to a point source above an homogeneous impedance plane and

$$P_1 = P_0 \frac{\exp(ikR_1)}{R_1} \tag{3}$$

is the pressure due to the direct wave from the source [using $\exp(-i\omega t)$ for the time dependence]. The variable Q is the spherical wave reflection coefficient, R_1 is the direct path length from source to receiver, R_2 is the path length through the specular reflection point (see Fig. 1), k is the wave number, and P_0 is a constant. De Jong's calculations give the following result for the pressure above a plane containing a single impedance discontinuity,
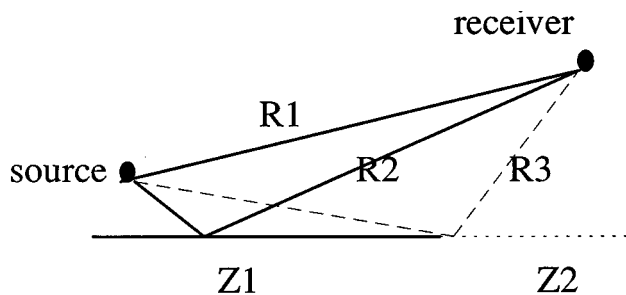
FIG. 1. Definition of the various path lengths.

$$\frac{P}{P_1} = 1 + \frac{R_1}{R_2} Q_{1,2} \exp\{ik(R_2 - R_1)\}$$

$$+ (Q_2 - Q_1) \frac{\exp(-i\pi/4)}{\sqrt{\pi}} \frac{R_1}{R_3}$$

$$\times \{F_{31} \pm F_{32} \exp\{ik(R_2 - R_1)\}\}, \tag{4}$$

where R_3 is the path length through the impedance discontinuity and Q_1 is the spherical wave reflection coefficient for the ground of impedance Z_1 used in conjunction with the plus sign in the right-hand side bracket when the specular point falls on the ground of impedance Z_1. If the specular point falls on the ground of impedance Z_2, the spherical wave reflection coefficient Q_2 is used in conjunction with the minus sign. In Eq. (4),

$$F_{31} \equiv F[\sqrt{k(R_3 - R_1)}] \tag{5}$$

and

$$F_{32} \equiv F[\sqrt{k(R_3 - R_2)}], \tag{6}$$

where

$$F(x) \equiv \int_x^{\infty} \exp(iw^2) dw \tag{7}$$

is the Fresnel integral function.

For convenience in the measurements, the source and receiver are placed at 0.1 m above the ground, and 1 m apart. The discontinuity between a finite impedance and a hard surface is perpendicular to the source–receiver line. The source is situated above the hard surface, and the receiver above the sand surface.

The admittance of the hard surface is modeled by choosing a value equal to zero, and the sand is modeled using a rigid porous medium model.[17] The model requires knowledge of flow resistivity, porosity, tortuosity, and the standard deviation of the (assumed) log normal pore size distribution. The flow resistivity of the sand used in the experiments was measured and the value was found to be 473 000 mks rayls/m. A common value of porosity for sand (0.4) and a tortuosity equal to the inverse of the porosity were assumed. A value of 1.3 for the standard deviation was found to provide the best visual fit between predictions and measurements of excess attenuation above sand.

2. The boundary element model

Chandler Wilde and Hothersall[14] have developed a boundary integral equation method to evaluate the pressure P at the receiver. A solution for the Helmholtz equation for the pressure at the receiver can be solved for a locally reacting boundary and gives, for the case of a strip with impedance Z_2 imbedded in a plane of impedance Z_1,

$$P(\mathbf{r}, \mathbf{r}_0) = P_1(\mathbf{r}, \mathbf{r}_0) - ik\left(\frac{1}{Z_2} - \frac{1}{Z_1}\right)$$

$$\times \int_s G_{Z_2}(\mathbf{r}_s, \mathbf{r}) P(\mathbf{r}_s, \mathbf{r}_0) ds(\mathbf{r}_s), \tag{8}$$

where $P_1(\mathbf{r}, \mathbf{r}_0)$ would be the pressure at the receiver if the surface had the homogeneous impedance Z_1 and $\mathbf{r}_0$, $\mathbf{r}$, $\mathbf{r}_s$ are the position vectors of the source, the receiver and a point in the boundary, respectively. $G_{Z_2}(\mathbf{r}_s, \mathbf{r})$ is the Green's function associated with propagation over a boundary of impedance Z_2, and s is the surface of the strip. To evaluate the integral, $P(\mathbf{r}_s, \mathbf{r}_0)$ is required over the whole strip and is calculated numerically by a standard boundary element technique. An equivalent two-dimensional problem can be solved in order to save computation time. Source, receiver, and specular point are assumed to be in a vertical plane perpendicular to the impedance discontinuity, and a line integral is solved instead of a surface integral. The boundary element code is applied easily to predicting propagation of sound over multiple strips by including more discontinuity coordinates in the input file to the program. The boundary element code used here was a modified version of that provided by the authors of Ref. 14 through a joint research program financed by an EPSRC (UK) grant. The original code was modified to allow for the impedance model described. The resulting predictions of the excess attenuation spectrum over a continuous sand surface are within 2 dB of the measured data for all measurement frequencies in the range 300 to 20 000 Hz.

3. The Fresnel-zone model

In order to evaluate the proportions of impedance of each type contributing significantly to the excess attenuation at the receiver, Slutsky and Bertoni[18] suggested use of Fresnel diffraction theory. This method has been developed and applied by Hothersall and Harriott.[19]

Consider points P' (see Fig. 2) such that

$$S'P' + RP' - R_2 = F\lambda, \tag{9}$$

where F is a constant, S is the source, R is the receiver, S' is the image source, and λ is the wavelength. Here, P' are points with a path length $1/3\lambda$ greater than the specularly reflected path; hence subtraction of the distance R_2 in Eq. (9). This consideration is based on the generalization to all space of the definition of Fresnel zones as regions in a surface over which ray paths from the source via reflection points to the receiver have path length in the range $n\lambda$ to $(n+1/2)\lambda$, where n is an integer. Hothersall and Harriott[11] have shown that path length differences between $\lambda/2$ and $\lambda/4$ are acceptable. Consequently, in the following work the value $F = 1/3$ will be used. The set of all the possible posi-
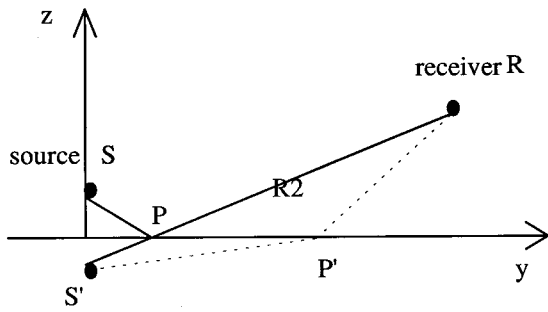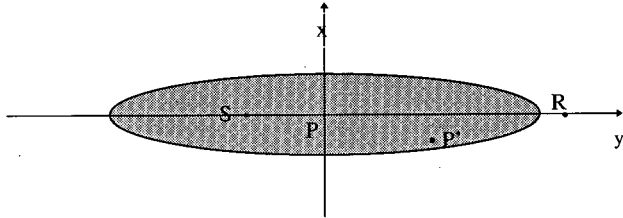
FIG. 2. Source receiver geometry (a) in the yz plane and (b) in the xy plane. (b) shows also the elliptical Fresnel zone, intersection of the horizontal plane, and the ellipsoid defined by Eq. (14).

tions P' defines an ellipsoid with foci S' and R. The major semiaxis of the ellipsoid along the line $S'R$ is given by

$$a = \frac{R_2 + F\lambda}{2} \tag{10}$$

and the minor semiaxis by

$$b = \sqrt{\frac{R_2 F\lambda}{2} + \left(\frac{F\lambda}{2}\right)^2}. \tag{11}$$

The position of the points P' relevant to our study in terms of Fresnel zones are inside the boundary, the goal being to assess the positions of the reflection points P' contributing significantly to the excess attenuation at the receiver. In order to do so, we need the intersection of the ellipsoid and the boundary. The form of this ellipse depends on the value of the constant F, the wavelength considered, and the source–receiver geometry. The equation of the ellipse in the coordinate system shown in Fig. 3 is

$$\frac{x^2}{b^2} + \frac{(y\cos\theta - c)^2}{a^2} + \frac{y^2 \sin^2\theta}{b^2} = 1, \tag{12}$$

where $c = R_2/2 - SP$. It should be noted that, in the equivalent equation in Ref. 19, the x and y variables are interchanged. This is inconsistent with the notation in that reference which is reproduced in Fig. 3. To compute the area of the ellipse we are interested in the boundary values $x_{1,2}$ and $y_{1,2}$ (see Fig. 3) and find

$$x_{1,2} = \pm b\sqrt{1 - \frac{(y_m\cos\theta - c)^2}{a^2} + \frac{y_m^2\sin^2\theta}{b^2}}, \tag{13}$$

where

$$y_m = \frac{B}{A} \tag{14}$$

and

$$y_{1,2} = -\frac{B}{A} \pm \sqrt{\frac{1}{A} - \left(\frac{c\sin\theta}{Aab}\right)^2}, \tag{15}$$

where $A = (\cos\theta/a)^2 + (\sin\theta/b)^2$ and $B = c\cos\theta/a^2$. It should be noted that the expression for $y_{1,2}$ given by Hothersall and Harriott[19] does not have the dimensions of length.

If we choose identical heights for the source and receiver, the specular reflection point is half way between source and receiver and $y_1 = y_2$. The ratio of the elliptical Fresnel zone surface area $\pi x_2 \times y_2$ and the rectangular area $d \times d/2$ are computed as a function of frequency for angles near grazing and for various source–receiver separation distances d. The width of the rectangle representing the propagation area is chosen to be an arbitrary fraction ($d/2$) of the separation distance d in order to compare better with the expected elliptical area. The results shown in Fig. 4 give an estimate of the proportion of rectangular surface area between source and receiver that is occupied by the elliptical Fresnel zone. Figure 4 shows the values of this ratio for frequencies in the range from (a) 10^2 to 10^3 Hz and (b) 10^3 to 10^4 Hz. The solid line, dotted line, dashed line, and dotted-dashed line represent results where the separation distances are 1, 2, 5, and 10 m, respectively. As an example, Fig. 4(a) shows that at 200 Hz the Fresnel zone surface area is three times greater than the surface area of the rectangle between source and receiver.

The first modified-Fresnel-zone model used in this study corresponds to that suggested by Hothersall and Harriott.[11,19] Following their assumption, we evaluate the excess attenuation at the receiver for a given geometry from

$$EA = \mu 20\log\left|1 + \frac{R_1}{R_2} Q_1 \exp\{ik(R_2 - R_1)\}\right| + (1 - \mu)$$
$$\times 20\log\left|1 + \frac{R_1}{R_2} Q_2 \exp\{ik(R_2 - R_1)\}\right|. \tag{16}$$

In Ref. 11 Hothersall and Harriott define μ as the proportion of the line representing the Fresnel zone and intersecting with the area of impedance Z_1. The line representing the Fresnel zone is defined as the intersection between the elliptical Fresnel zone and the vertical plane containing source,
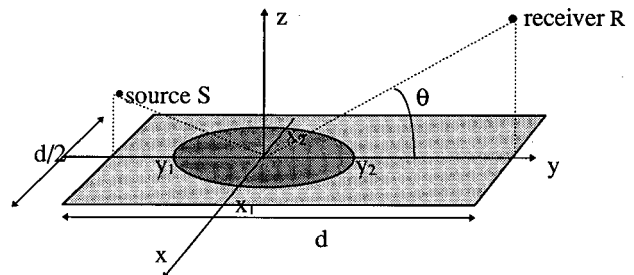
FIG. 3. Three-dimensional view of the source–receiver geometry and the elliptical Fresnel zone in the boundary.
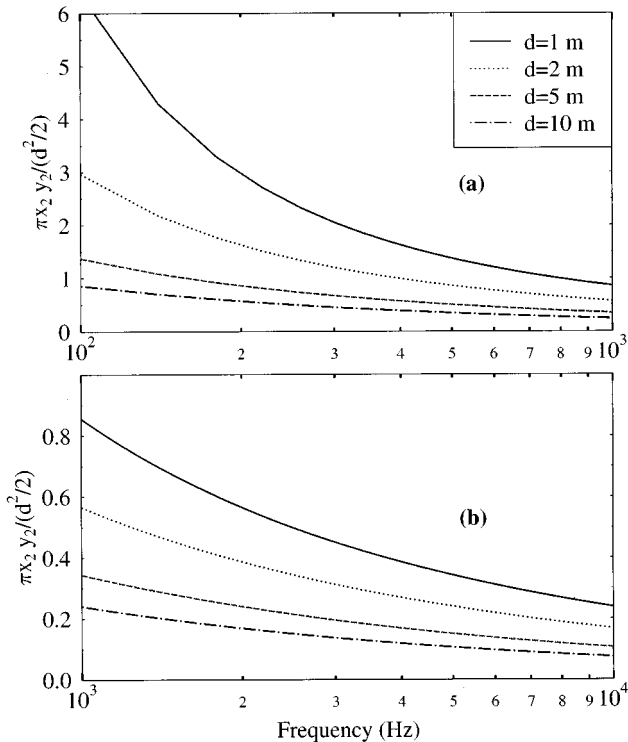
FIG. 4. Ratio of the Fresnel zone surface area $\pi x_2 y_2$ and the ground surface area $d^2/2$ for frequencies in the range from (a) 10^2 to 10^3 Hz and (b) 1 to 10 kHz.

FIG. 5. Definition of the various path lengths in the case of a single strip.

receiver, and specular reflection point. The first term in Eq. (16) is the product of μ and the excess attenuation EA_1 above a uniform boundary of surface impedance Z_1, and the second term is the product of $(1-\mu)$ and the excess attenuation EA_2 above a uniform boundary of surface impedance Z_2. On the other hand Hothersall and Harriott[11] give a definition of the excess attenuation that is inconsistent with Eq. (16). In the limits $\mu=0$ or $\mu=1$ the excess attenuation defined in Ref. 11 does not tend to the expected values of EA_2 or EA_1, respectively. Nevertheless, since the graphs of their predicted excess attenuation versus plan distance from the source to the impedance discontinuity exhibit straight lines, as is expected from a linear interpolation between the limits EA_2 and EA_1, it is clear that they used Eq. (16) for their calculations of the excess attenuation. This conclusion is supported by Ref. 19.

An alternative is to assume that the excess attenuation depends linearly on the proportions of the different surface areas inside the elliptical Fresnel zone in the plane of the boundary.[19] The coefficient μ needed to compute the excess attenuation from relation (16) could be obtained by calculating the proportion of area having surface impedance Z_1 inside the elliptical Fresnel zone. However, greater improvement in the model can be achieved by modifying the definition of the excess attenuation from that used in Eq. (16).

The second modified-Fresnel-zone model proposed in this work is based on a new definition of the excess attenuation in mixed impedance cases. Instead of the linear interpolation between EA_1 and EA_2 shown in relation (16), we use a linear interpolation between two pressure terms. The

first term is the pressure at the receiver relative to the free field in the case of a surface of impedance Z_1 and the second term is the pressure at the receiver relative to the free field in the case of a surface of impedance Z_2. The new expression proposed for the excess attenuation is

$$EA = 20 \log\left\{ \mu \left| 1 + \frac{R_1}{R_2} Q_1 \exp\{ik(R_2-R_1)\}\right| + (1-\mu)\right.$$

$$\left. \times \left| 1 + \frac{R_1}{R_2} Q_2 \exp\{ik(R_2-R_1)\}\right| \right\}. \tag{17}$$

As previously, the coefficient μ is obtained by calculating the proportion of area having surface impedance Z_1 inside the elliptical Fresnel zone.

We note that the Fresnel-zone model can be extended easily to multiple strips, but, as can be seen in the definition of the excess attenuation in relation (17), for a given ratio of hard and soft surface, it does not differentiate between situations with either a single discontinuity or multiple strips.

B. Theories of sound propagation over a periodic impedance

1. Extension of the De Jong model to a periodic impedance

It has been shown in earlier work by Hothersall and Harriott[11] that it is straightforward to apply the empirical model developed by De Jong to a single strip. However, in their expression for the pressure at the receiver relative to the free field [their Eqs. (10), (11), and (12)], the Fresnel integral $F[\sqrt{k(R_5-R_1)}]$ should be multiplied by -1 to yield the correct limit for a single discontinuity when the position of one of the two discontinuities defining the strip tends to infinity. Figure 5 defines the various path lengths needed in the expression of the acoustic pressure over a ground of impedance Z_2 with a single strip of impedance Z_1. No strip outside the area between the source and the receiver is considered because it is assumed that the strips between the source and the receiver contribute a major part at frequencies above 1000 Hz.

It is straightforward to extend the model to multiple strips by summing over all the diffraction terms, each term arising from each impedance discontinuity. This method is outlined by Bassiouni et al.[20] and is detailed in the follow-
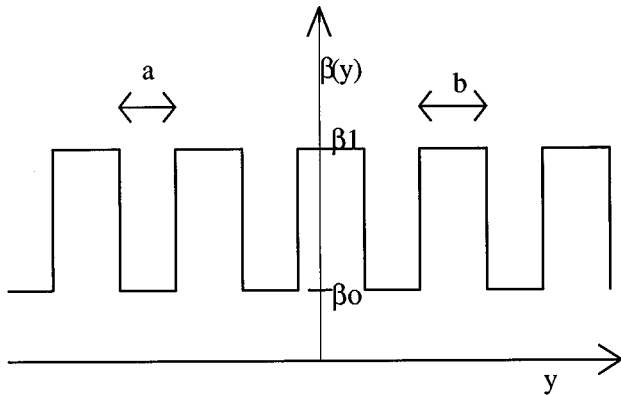
FIG. 6. The periodic admittance used in Nyberg's theory.

ing. The pressure at the receiver relative to the free field in the case of multiple strips of impedance Z_1 imbedded in a plane of impedance Z_2 is given by

$$\frac{P}{P_1} = 1 + \frac{R_1}{R_2} Q_{1,2} \exp\{ik(R_2 - R_1)\}$$

$$+ (Q_2 - Q_1) \frac{\exp(-i\pi/4)}{\sqrt{\pi}} R_1$$

$$\times \sum_{j=1}^{n} \left[\frac{1}{R_{b_j}} \{F_{b_j1} + \gamma_j F_{b_j2} \exp\{ik(R_2 - R_1)\}\} \right.$$

$$\left. + \frac{1}{R_{aj}} \{-F_{a_j1} + \delta_j F_{a_j2} \exp\{ik(R_2 - R_1)\}\} \right], \quad (18)$$

where j is the index characterizing the various strips and n is the total number of strips between source and receiver. In accordance with previous definitions, $F_{a_j1} \equiv F(\sqrt{k(R_{a_j} - R_1)})$. The spherical wave reflection coefficients Q_1 or Q_2 are used when the specular reflection point falls on the ground of impedance Z_1 or Z_2, respectively. The sequences of values $(Q_{1,2}, \gamma_j, \delta_j)$ appear in Eq. (18) according to the following combinations: $(Q_2, 1, -1)$ if the specular reflection point is on the source side of the jth strip, $(Q_2, -1, 1)$ if the specular reflection point is on the receiver side of the jth strip, $(Q_1, -1, -1)$ if the specular reflection point is inside the jth strip,

2. Nyberg's theory

In Ref. 13, Nyberg shows that the Helmholtz equation with the boundary condition for a point source above an infinite plane surface with striped impedance can be solved by using Cartesian coordinates and a Fourier transform technique. The ground effect due to a two-valued infinitely periodic striped impedance (such as that shown in Fig. 6) may be determined from that predicted by using the area-averaged impedance $(\beta_0 b + \beta_1 a)/(a + b)$, where a and b are the periods of the two types of strips. The approximation made in the derivation of the theory is represented by $a + b \ll \lambda$ for the particular geometry used in our experiment. This condi-

tion is satisfied for frequencies up to 1700 Hz when the proportion of hard surface is 50% and up to 3400 Hz in the case of 100% hard surface.

II. DESCRIPTION OF THE EXPERIMENTAL PROCEDURE

A. Experiments involving sound propagation over a single impedance discontinuity

The point source used in the experiment was a Tannoy driver fitted with a 1-m-long tube, with a 3-cm internal diameter. It was suspended 0.1 m above the surface of the sandbox in an anechoic chamber of dimensions $3 \times 3 \times 3$ m. The receiver was a Bruel & Kjaer type 4311 1.3-cm-diam condenser microphone fitted with a preamplifier suspended 1 m away at the same height. Signal processing and signal generating were carried out using MLSSA, a maximum length sequence system analyzer. Background noise effects were eliminated by analyzing the microphone signal with respect to the known output sequence. In order to calculate the excess attenuation, a reference measurement was made with source and receiver at the maximum height of 2 m allowed by the suspending cable above the floor of the anechoic chamber. A hardwood board measuring $1.1 \times 0.9 \times 0.03$ m was set flush with the surface of the sand in a sandbox measuring $2 \times 1.2 \times 0.3$ m.

In this experiment the proportion of hard surface was varied, while trying to minimize edge effects, either by moving the source and receiver normal to the discontinuity or by translating the board in the sand pit parallel to the source receiver axis.

B. Experiments involving sound propagation over a periodic impedance

The experimental procedure was very similar to the one previously used in measurements over a single impedance discontinuity and source and receiver heights and separation distances were kept identical. The main change was the use of a smaller sandbox of size $1.0 \times 0.8 \times 0.08$ m (deep) in which wooden blocks of size $1.0 \times 0.1 \times 0.08$ m were moved, inside the sand pit, along the source–receiver line perpendicularly to their axes in order to achieve the various proportions of hard surface.

III. COMPARISON BETWEEN EXPERIMENTAL RESULTS AND PREDICTIONS

A. Sound propagation over a single impedance discontinuity

In the work presented here and unless specified otherwise, for example, a 40% hard surface between source and receiver means that the discontinuity between sand and hard board is at a horizontal distance of 0.4 m from the source. For the special case 0% hard surface (discontinuity beneath the source) the measurements compare well with the case of a continuous sand surface and the latter will be implied when a 0% hard surface is mentioned. The excess attenuation spectra measured for proportions of hard surface 50%, 80% are shown as the dotted lines in Fig. 7 and Fig. 8, respectively.
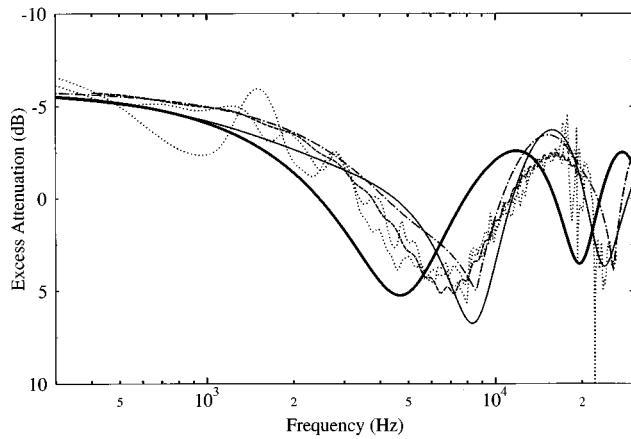
FIG. 7. Excess attenuation spectra for a 50% hard boundary. The dotted lines are the measured data, the thin solid line is the prediction using the De Jong model, the dashed line is the boundary element prediction, and the dashed-dotted line is the second modified Fresnel-zone model prediction. The thick solid line is the prediction using the De Jong model for a continuous sand boundary.

Figure 7 presents two measured spectra for the case of 50% hard surface to show the overall good reproducibility of the measurements. Figures 7 and 8 display also the predictions using the second modified-Fresnel-zone method shown as a dotted-dashed line, the predictions using the De Jong semi-empirical model shown as a thin solid line, and the predictions using the boundary element code shown as a dashed line. The excess attenuation spectrum predicted by the De Jong model reduces to the classical Weyl–van der Pol expression for a continuous sand surface, and is shown as the thick solid lines in Fig. 7. The overall agreement between measurements and predictions of the depths and frequency positions of the excess attenuation spectra first dip is satisfactory.

The frequency of the minima of the first dip in measured excess attenuation spectra (ground effect) is seen to increase steadily from the cases of low percentage of hard surface to the cases of high percentage up to 70%. However, for values of the proportion of hard surface above 70%, the frequencies
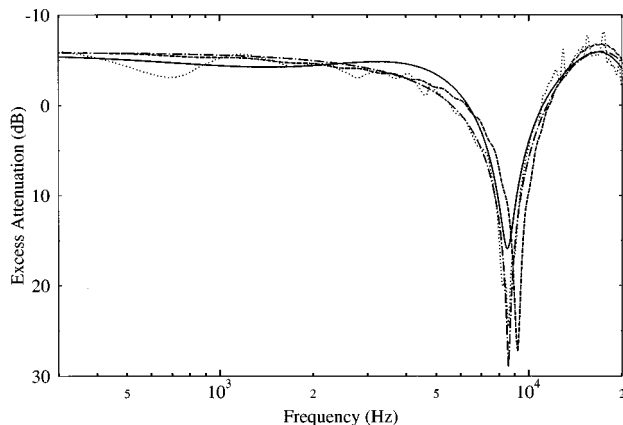
FIG. 8. Excess attenuation spectra for a 80% hard boundary. The dotted line is the measured data, the thin solid line is the prediction using the De Jong model, the dashed line is the boundary element prediction, and the dashed-dotted line is the second modified Fresnel-zone model prediction.
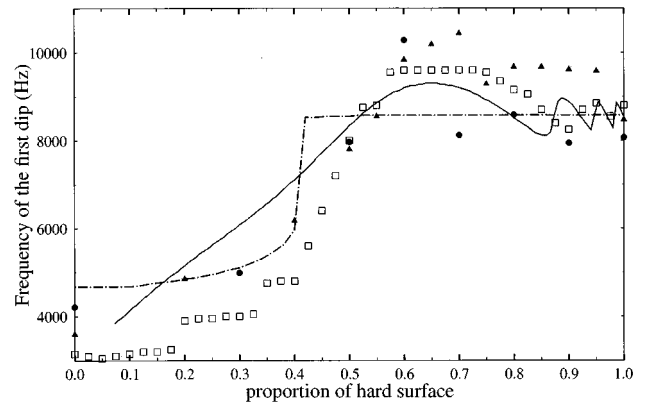
FIG. 9. Frequency location of the first dip for the excess attenuation spectra in the case of varying proportions of hard surface. The circles and triangles are two different data sets, the squares are the boundary element predictions, the solid line is the prediction using the De Jong model, and the dotted-dashed line is the prediction from the second modified-Fresnel-zone model.

of the minima of the first dip are seen to decrease as the percentage of hard surface approaches 100%. In order to show clearer evidence of this phenomenon, more experimental data have been taken and are shown in Fig. 9 where the frequencies of the first dip minima are plotted versus the proportion of hard surface. The circles are obtained in part from the data presented in Figs. 7 and 8, and the triangles are data taken in a second experiment involving more cases of various proportions of hard surface. The measurements show a lack of repeatability in dip frequency position for the cases involving 70%, 80%, and 90% of hard surface. However, this statement is related to the degree of fit in a region of the spectrum which is very sensitive to the geometry. The measured shift in dip frequency position can be reproduced in the predictions by a 7% variation in source and receiver heights. Therefore slight variations in the location of the reflecting plane might be the cause of the lack of repeatability. Nevertheless, among the 14 measurements taken for proportions of hard surface 60% and above, nine of them show a frequency position of the first dip above the expected maximum value of 8500 Hz. Figure 9 shows also the predictions from the improved Fresnel-zone method as a dotted-dashed line, the predictions from the De Jong semiempirical model as a solid line, and the predictions from the boundary element code as squares. The improved Fresnel-zone method shows a discontinuity in the predicted dip frequency when the horizontal distance from source to discontinuity is near to 0.4 m. This discontinuity is expected from the definition of the excess attenuation given in Eq. (17) which is a weighted sum of the excess attenuation spectra of a continuous sand surface and a continuous hard surface. However, this discontinuity is not displayed by the measurements and shows the limitation of the Fresnel zone method in predicting the exact frequency of the ground effect in excess attenuation spectra. The boundary element theory predicts a maximum dip frequency of 9600 Hz for proportions 60%–70% while the De Jong model predicts a dip frequency maximum of 9300 Hz at a hard surface percentage of 65%. The measured maximum dip frequency is 10 500 Hz and occurs at 70% hard surface. It is clear from Fig. 10(a) and (b) that the existence of this maximum is the

FIG. 11. Excess attenuation spectra for a 70% hard ground cover. The dotted line is the measured data, the solid line is the prediction using the first modified Fresnel-zone model, the dashed line is the prediction using the crude model, and the dashed-dotted line is the prediction using the second modified-Fresnel-zone model.

FIG. 10. Excess attenuation spectra for (a) a 65% hard ground cover, (b) a 80% hard ground cover. The solid line is the prediction $20 \log|1+A+D|$ using the De Jong model and displayed in the envelope defined by $20 \log(|1+A|-|D|)$ and $20 \log(|1+A|+|D|)$. $1+A$ is the Weyl–van der Pol term and D is the diffraction term.

result of diffraction of the sound wave by the single discontinuity. In Fig. 10(a) and (b) the dotted line, dashed line, and the solid line display the function $20 \log(|1+A|-|D|)$, $20 \log(|1+A|+|D|)$, and the excess attenuation spectrum $20 \log|1+A+D|$, respectively. The term $1+A$ stands for the Weyl–van der Pol term [first two terms in Eq. (4)] and D stands for the diffraction term [last term in Eq. (4)]. Figure 10(a) shows the results in the case 65% of hard surface, where the predicted first dip frequency is highest. Figure 10(b) shows the results in the case 80% of hard surface. We note that for the phenomenon to occur, the terms A and D need to have opposite phases for frequencies in the higher portion of the frequency range where $1+A$ and D have comparable magnitudes (in this case $20 \log(|1+A|-|D|)$ reaches it's minimum region). Secondary maxima are observed in Fig. 9 for proportions of hard surface 89% and 95%. They are explained by the phase of the diffraction term varying more rapidly as the proportion of hard surface increases, and reaching a maximum rate of variation when the discontinuity is situated vertically beneath the receiver. Consequences are the presence of oscillations in the excess attenuation spectrum when the receiver is close to the single discontinuity.

This phenomenon is suggested by Fig. 3 in Ref. 14 where measured excess attenuation spectra are plotted for a

source height of 85 mm, a receiver height of 53 mm, and a separation distance of 1 m. In this study the frequency of the minima of the first dip (ground effect) was seen to increase for the proportions of hard surface 0%, 25%, and 50%. However, the shape of the graph for the highest frequencies studied in the 75% case suggests that the frequency of the minima of the first dip is higher than the frequency of the minima for the 100% case. A study presented by Klein[21] shows the same predictions for a source height of 1.83 m, a receiver height of 1.5 m, and a separation distance of 62 m. The model used in this study was based on a numerical integration and the excess attenuation functions predicted in Fig. IV.3 showed clearly that, when the source is above the hard surface, the ground effect in the 75% hard surface case occurs at higher frequency than for the continuous hard surface. Also, when the source is above the soft surface, the high-frequency behavior of the graph for the 75% hard surface case suggests a ground effect at higher frequency than for the continuous hard surface (100%) case. In neither Ref. 14 nor Ref. 21 is the exact location of the minima for the cases 75% hard surface exhibited since the range of frequencies studied is too narrow.

Prediction of the excess attenuation spectra from the first modified- and second modified-Fresnel-zone model for a proportion of hard surface equal to 70% are shown as the solid line and the dotted-dashed line, respectively, in Fig. 11. The agreement between measurements and the prediction from the first modification of the Fresnel-zone model [Eq. (16)] is poor. Indeed, the magnitude of the dip is over predicted by 20 dB. These results demonstrate that the first modification of the Fresnel-zone method is not satisfactory for predicting these laboratory measurements of excess attenuation spectra. On the other hand, the agreement between the measured data and the predictions of the second modification [Eq. (17)] is good. Figure 11 also establishes clearly the improvement of the predictions due to the introduction in Eq. (17) of a frequency-dependent coefficient μ (as in the Fresnel-zone model) compared with predictions from a crude model (dashed line) where the excess attenuation is also
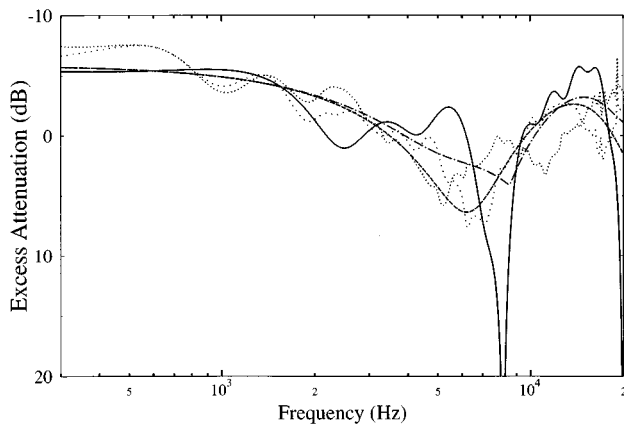
FIG. 12. Excess attenuation spectra for a 50% hard-strip boundary. The closely packed dotted line is the measured data in the case of a strip of sand at the specular reflection point, the other dotted line is the measured data for a strip of wood at the specular reflection point, the solid line is the prediction using the De Jong model, the dashed line is the prediction using Nyberg's theory, and the dashed-dotted line is the second modified-Fresnel-zone model prediction.

given by Eq. (17) but the coefficients μ are frequency independent. For this crude model, a horizontal distance 0.7 m between source and discontinuity implies that 70% of the surface is covered with hard ground and thus $\mu = 0.7$ for all frequencies. The crude model allows good predictions of the frequency position of the first dip, but the second modification of the Fresnel-zone model allows much better prediction of the magnitude of the dip.

B. Sound propagation over a periodic impedance

The measured excess attenuation spectra for proportions of hard surface 50%, 70%, and 80% are shown as the dotted lines in Figs. 12 and 13(b). In the case of 50% of hard surface, Fig. 12 shows two measurements. The dotted line with closely packed dots corresponds to a sand strip at the specular reflection point, and the other represents the case when a hard strip is at the specular reflection point. The two measured excess attenuation curves are slightly shifted with an increase in frequency of the minimum of the first dip in the case of the specular reflection point on the strip of wood as would be expected. The predictions obtained from Eq. (17) and from Nyberg's theory are shown as the dotted-dashed lines and dashed lines, respectively. In Fig. 12 the solid line is the prediction using the extended De Jong model and in Fig. 13(a) and (b) the solid line is the prediction using the boundary element calculations.

The measured data show oscillations that could be mistaken for diffraction effects due to the array of strips. However, such oscillations appear also in data obtained over continuous sand. These fluctuations in excess attenuation appear to be due to the reflections of sound at the edges of the sandbox used for the experiments. These fluctuations probably generate inaccuracies in the measurement of the frequency of the first minima in the spectra.

A typical result of the extended De Jong model prediction is shown in Fig. 12 and displays clear discrepancies when compared to the measurements. The depth of the first
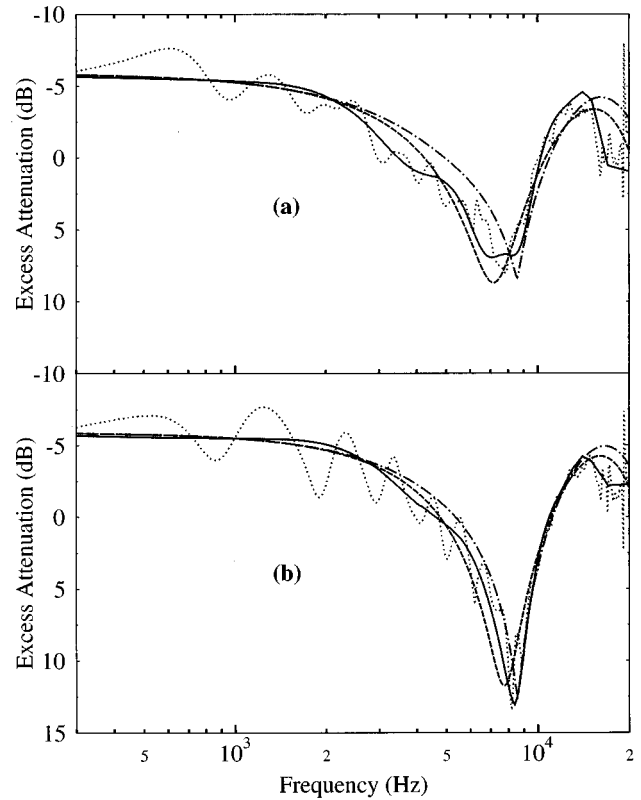
FIG. 13. Excess attenuation spectra for (a) a 70% hard-strip boundary (a strip of wood at the specular reflection point), (b) a 80% hard-strip boundary (a strip of sand at the specular reflection point). The dotted line is the measured data, the solid line is the boundary element prediction, the dashed line is the prediction using Nyberg's theory, and the dashed-dotted line is the second modified-Fresnel-zone model prediction.

dip is over predicted by the model by more than 20 dB. The same type of discrepancy is observed for the other cases of strip proportion, and therefore, predictions from the extended De Jong model are not reported further.

In Fig. 13(a) and (b) the agreement between measured data and the predictions of Nyberg's theory is satisfactory and appears even better where there is a higher percentage of hard surface. The average-admittance form of Nyberg's theory cannot distinguish between the situation where one type of strip or another is at the specular reflection point for a given percentage of hard surface. Therefore only one prediction is obtained for the 50% case. However, the prediction seems to be a good average of the two measured cases for either a wood or sand strip at the specular reflection point. For all cases the agreement appears to be surprisingly good for values of frequency above 2000 Hz even though the condition $a+b \ll \lambda$ might not be satisfied. Further analysis of the quasiperiod approximation obtained by Nyberg[13] might be needed to explore whether the approximation is valid for higher frequencies or wider strips than satisfy the condition $a+b \ll \lambda$.

Figure 13(a) and (b) show very good agreement between the boundary element code predictions and the measured data. Figure 14(a) and (b) present results for a 60% hard surface when a strip of wood or a strip of sand are at the specular reflection point, respectively. The very good agreement between the predictions and the measured data show
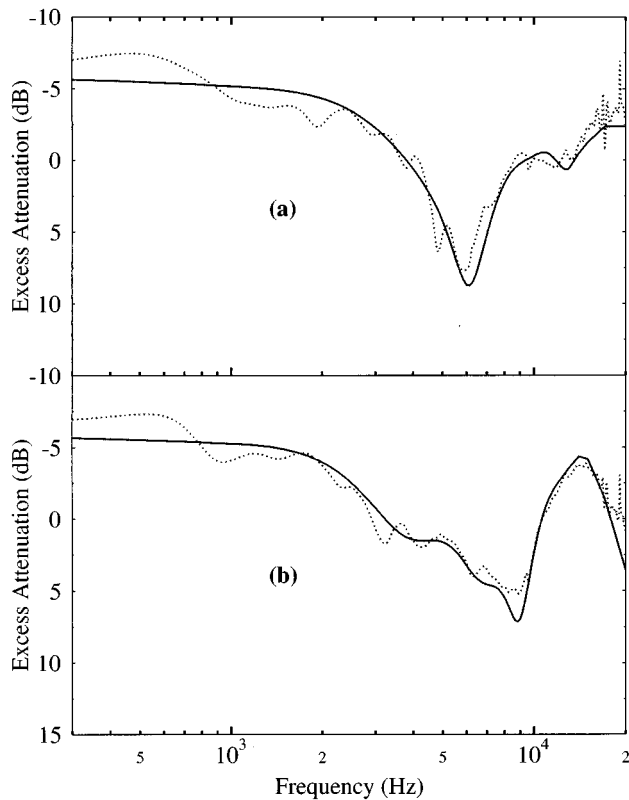
FIG. 14. Excess attenuation spectra for a 60% hard-strip boundary with (a) a strip of wood at the specular reflection point, (b) a strip of sand at the specular reflection point. The dotted line is the measured data and the solid line is the boundary element prediction.

that the boundary element code can distinguish between cases where a hard strip or a sand strip is at the specular reflection point for a same percentage of hard strip, and can predict the excess attenuation spectra adequately. None of the other models studied in this work are able to distinguish between such cases.

Figures 12 to 13(b) show a qualitative agreement between the measured data and the predictions of the second modified-Fresnel-zone model. The magnitude of the first dip is predicted correctly but discrepancies are found for the frequency position of the dip as the proportion of hard strips decreases.

IV. CONCLUSIONS

Laboratory measurements of the excess attenuation of sound from a point source over a single discontinuity between an acoustically hard surface and a sand surface have been compared to (a) a De Jong semiempirical model, (b) a Fresnel-zone model, and (c) a boundary element theory when the proportion of hard surface is varied systematically. The Fresnel-zone method used in our work is improved compared to previous models which derived incorrect boundary values of the elliptical Fresnel zone. A second improvement is the use of a new definition of the excess attenuation in mixed impedance cases. Instead of a linear interpolation between the predicted excess attenuation above uniform boundaries of each impedance, we use a linear interpolation between the pressure at the receiver relative to the free field

above a surface of impedance Z_1 and the pressure at the receiver relative to the free field above a surface of impedance Z_2. The results of the predictions obtained by this modified Fresnel-zone model show a better agreement with the measured data presented here than that calculated from the previous Fresnel-zone model. The De Jong semiempirical model and the boundary element code are found to give satisfactory agreement with the measurements but the modified Fresnel-zone model does not perform as well in predicting the frequency position of the first dip. In addition, due to diffraction effect, the frequency of the excess attenuation first dip is predicted to be highest at approximately 70% hard surface cover rather than the expected 100%.

Laboratory measurements of the excess attenuation of sound from a point source over a striped impedance with the source–receiver line perpendicular to the strips are compared to (a) the De Jong semiempirical model extended to the case of multiple impedance discontinuities, (b) to the modified Fresnel-zone model, (c) to the boundary element method, and (d) to Nyberg's theory. The results show that the De Jong semiempirical model and the modified Fresnel-zone model are not adequate to predict excess attenuation for all frequencies in the case of multiple impedance discontinuities. However, both Nyberg's theory and the boundary element method give good agreement with the measured results. The agreement between Nyberg's theory and the measurements shows that for the geometry used in the experiment, the ground effect due to striped impedance may be determined from that predicted by using the area-averaged impedance.

[1] M. Naghieh and S. Hayek, ''Diffraction of a point source by two impedance covered half-planes,'' J. Acoust. Soc. Am. **69**, 629–637 (1981).

[2] B. O. Enflo and P. H. Enflo, ''Sound wave propagation from a point source over a homogeneous surface and over a surface with an impedance discontinuity,'' J. Acoust. Soc. Am. **82**, 2123–2134 (1987).

[3] J. Naze Tjøtta, ''Reflection of an acoustic wave on a plane surface presenting an impedance discontinuity along a straight line,'' Acta Acust. (China) **1**, 15–33 (1993).

[4] J. Durnin and H. Bertoni, ''Acoustic propagation over ground having inhomogeneous surface impedance,'' J. Acoust. Soc. Am. **70**, 852–859 (1981).

[5] J. N. B. Harriott, S. N. Chandler-Wilde, and D. C. Hothersall, ''Long distance sound propagation over an impedance discontinuity,'' J. Sound Vib. **148**, 365–380 (1991).

[6] K. B. Rasmussen, ''Propagation of road traffic noise over level terrain,'' J. Sound Vib. **82**, 51–61 (1982).

[7] M. Galindo, ''Application of the parabolic approximation method to sound propagation above ground with impedance variations,'' *Proceedings of the Sixth International Symposium on Long-Range Sound Propagation* (National Research Council, Canada, 1994), pp. 394–407.

[8] J. S. Robertson, P. J. Schlatter, and W. L. Siegmann, ''Sound propagation over impedance discontinuities with the parabolic approximation,'' J. Acoust. Soc. Am. **99**, 761–767 (1996).

[9] P. Koers, ''Diffraction by an absorbing barrier or by an impedance transition,'' Proc. Internoise, **1**, 311–314 (1983).

[10] G. A. Daigle, J. Nicolas, and J.-L. Berry, ''Propagation of noise above ground having an impedance discontinuity,'' J. Acoust. Soc. Am. **77**, 127–138 (1985).

[11] D. C. Hothersall and J. N. B. Harriott, ''Approximate models for sound

propagation above multi-impedance plane boundaries,'' J. Acoust. Soc. Am. **97**, 918–926 (1995).

[12] B. A. De Jong, A. Moerkerken, and J. D. Van Der Toorn, ''Propagation of sound over grassland and over an earth barrier,'' J. Sound Vib. **86**, 23–46 (1983).

[13] C. Nyberg, ''The sound field from a point source above a striped impedance boundary,'' Acta Acust. (China) **3**, 315–322 (1995).

[14] S. N. Chandler-Wilde and D. C. Hothersall, ''Sound propagation above an inhomogeneous impedance plane,'' J. Sound Vib. **98**, 475–491 (1985).

[15] K. Attenborough, ''Models for the acoustical properties of air-saturated granular media,'' Acta Acust. (China) **1**, 213–226 (1993).

[16] S. N. Chandler-Wilde and K. V. Horoshenkov, ''Padé approximants for the acoustical characteristics of rigid frame porous media,'' J. Acoust. Soc. Am. **98**, 1119–1129 (1995).

[17] K. Attenborough, ''Models for the acoustical properties of rigid-porous materials,'' Proceedings of the 15th International Congress of Acoustics **2**, 365–368 (1995).

[18] S. Slutsky and H. L. Bertoni, ''Analysis and programs for assessment of absorptive and tilted parallel barriers,'' Transportation Research Record 1176, Transportation Research Record, National Research Council, Washington, DC, 13–22 (1987).

[19] D. C. Hothersall and J. N. B. Harriott, ''A Fresnel zone approach to the prediction of sound propagation above a multi-impedance plane,'' Proc. Inst. Acoust. **16**, 83–90 (1994).

[20] M. R. Bassiouni, C. R. Minassian, and B. Chang, ''Prediction and experimental verification of far field sound propagation over varying ground surfaces,'' Proc. Internoise, **1**, 287–290 (1983).

[21] C. Klein, ''Kritische studie van het geluidsveld, opgewekt door een monopool, in de nabijheid van een reflekterend oppervlak'' Ph.D. dissertation, Catholic University of Leuven, 1983.

Effects of strong sound velocity gradients on propagation of low-frequency impulse sound: Comparison of fast field program predictions and experimental data

L. R. Hole[a)]
Norwegian Defence Construction Service (NDCS), Oslo, Norway

P. Lunde
Chrstian Michelsen Research AS (CMR), Fantoftvegen 38, P.O. Box 3, N-5036 Fantoft, Bergen, Norway

Y. T. Gjessing
University Courses on Svalbard (UNIS), P.O. Box 156, N-9170 Longyearbyen, Norway

As part of a large outdoor sound propagation experiment series, propagation of low-frequency impluse sound at ranges from 100 to 1400 m was studied at Haslemoen, Norway in February 1995. Sound sources were charges of 1 and 8 kg C4 explosives. Experiments were carried out in both a uniform forest and above a flat and uniform open field. Extensive meteorological measurements were carried out. Both automatic weather stations mounted in permanent towers and a tethered balloon were employed. The measurements resulted in a manifold set of sound velocity profiles. Occasionally, profiles were pure downwind or upwind, but mostly they were complex, often with large vertical gradients. One conclusion is that one should not uncritically use simple profile parameterizations, e.g., logarithmic, in sound propagation models. A fast field program (FFP), CAPROS, handles the complex sound-speed profiles well, predicting propagation of sound at 8, 16, 32.5, and 63 Hz. In this model, ground is considered a viscoelastic medium. At 63 Hz, the forest seems to cause an excess attenuation. At the lower frequencies, distinct effects of forest are not observed. There is no evidence that FFP predictions are less accurate in periods with high turbulence levels, which is here characterized with dynamical instability. However, variability of meteorological profiles causes predictions to be less accurate. This investigation suggests that it is sufficient to sample atmospheric variables up to 300 m above ground to predict sound propagation at the ranges studied here. © *1997 Acoustical Society of America.* [S0001-4966(97)03108-1]

PACS numbers: 43.28.Fp, 43.28.Dm, 43.20.Bi [LCS]

INTRODUCTION

Several papers have compared predictions by the wave number integration technique, or fast field programs (FFP), with experimental data of atmospheric sound propagation.[1–4] In addition, some publications have dealt with theoretical verification and modification of models of this type and comparison with other types of models.[5–9] FFP models have proven to give satisfactory results for a variety of propagation conditions and for sources of different nature. Low-frequency sound propagation has been studied by some authors, e.g., Robertson *et al.* (1995)[10] who presented parabolic approximation predictions of pure-tone frequencies from 40 to 160 Hz out to 3 km and compared with experimental data.[11] An extensive experimental study of propagation of wind turbine noise at the 10-Hz one-third octave band out to 20 km has also been carried out together with a theoretical analysis.[12,13] However, as far as we know, there have not yet been published any results of comparison of experiments and FFP predictions of propagation of impulse, infrasound frequencies at the ranges studied here (out to 1400 m). To give reliable predictions of propagation of sound in the frequency range studied here is of practical interest, since atmospheric

absorption is weak, about 0.1 to 0.5 dB/km.[14] Thus one can expect relatively low sound propagation losses at ranges where meteorological effects are significant. The relative importance of different effects on atmospheric sound propagation at different frequencies has recently been reviewed by Larsson (1994).[15]

Espérance *et al.* (1995)[3] showed that effects of atmospheric turbulence can be successfully predicted by the FFP code in some cases. This effect is important in many situations for higher frequencies.[15] In the present investigation, propagation of impulse sound from 1- and 8-kg charges of C4 explosives has been studied. In the linear region, which is of interest here, most of the energy from these charges is contained below 100 Hz. Impulse sound in this frequency range can be annoying, partly because it might cause buildings to shake. Since we expect that this sound is less affected by turbulence, it is adequate to base predictions of sound propagation on mean sound velocity profiles, Thus when predicting sound levels along ground, it is most important to consider atmospheric refraction and ground interaction. For sound propagation in this frequency range, above frozen ground, with strong atmospheric effects as studied here, we assume ground porosity is of secondary importance.[16] Consequently, a viscoelastic approach was chosen. In this paper,

---
[a)]Presently at UNIS. Electronic mail: larshole@unis.no
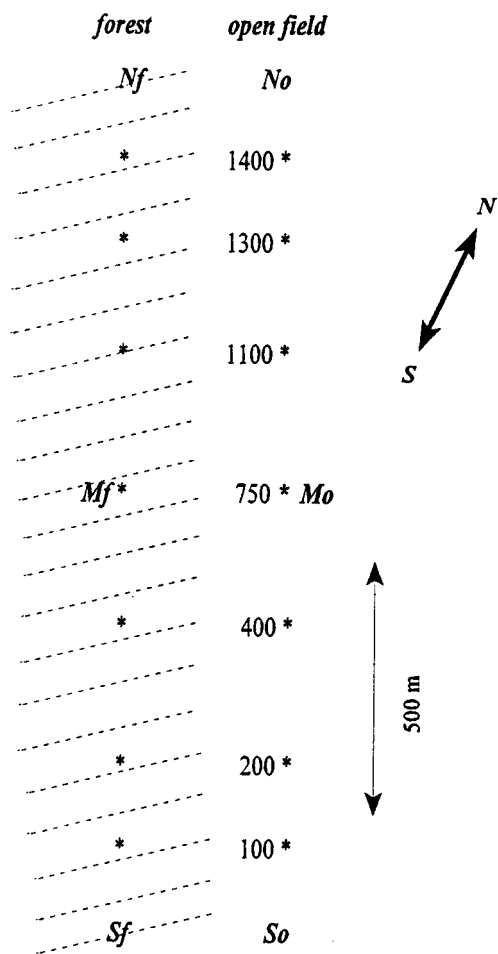
forest   open field

FIG. 1. Plan view of experiment layout. Shot positions are shown as asterisks (*). Numbers are distance in meters from south tower. Nf: North forest tower. No: North open field tower Sf: South forest tower. So: South open field tower. Automatic Weather Stations were placed at: Mf: Forest meteorology tower, Mo: Open field meteorology tower. Wind direction was sampled at 30 m at Nf. The scale is indicated on the right.

propagation of sound at ranges from 100 to 1400 m above uniform ground is investigated.

The data presented here were sampled as part of an international experimental sound propagation series called *Norwegian Trials*. All data from these trials will be incorporated in an open database that is being established. The database will contain well-documented data for propagation of low-frequency, impulse sound at ranges from 100 to 24 000 m, both in uniform and nonuniform terrain. Meteorological, seismic, topographical, and ground impedance data and will be available as well. The database will be completed in 1998, and can then be exported both in CD-ROM format and via

Internet, as well as through the Physics Auxiliary Publication Service (PAPS) of the American Institute of Physics. About 50% of the data are already available.

I. DESCRIPTION OF SITE AND EXPERIMENT

The trial was undertaken on a flat and uniform open grass field and in a uniform pine forest from 20 to 23 February 1995 (Fig. 1). The site is close to the Haslemoen army camp in Hedmark county, about 100 km north of Oslo. Tree height was around 18–19 m and average distance between trees was approximately 5 m. Most of the data presented here are from the open field, but one shot series from the forest is also included for comparison. Parts of the collected data can be used to study effects of forest on sound propagation, but that is beyond the scope of the present investigation.

A 0.1-m-thick snow cover was present during the trial, and on 22 and 23 February the snow was wet. At the low frequencies considered here (8–63 Hz), with wave lengths in the range 5–40 m, we assume this thin snow cover will not significantly affect sound-pressure levels along the ground. A frozen crust was present at the ground surface. This crust was associated with large wave speeds. Ground data as provided by Norwegian Geotechnical Institute (NGI)[17] during the trial are present in Table I. These data were obtained by the spectral analysis of surface waves method.[18] The shear (S) wave speed is set to half the compressional (P) wave speed.[17] Test runs of a full wave simulation model described later, CAPROS, showed that layers below those shown here did not affect predicted sound levels along the ground. Bottom layer (No. 4) identified in Table I is included as a half-space in CAPROS. Gradients of the P-wave speed were smoother at greater depths. Thus only the four layers shown in Table I were retained in later runs, and computation time was reduced. Experimental layout and positions of instruments are indicated in Fig. 1. Shots 250–298 were detonated in the forest, while shots 299–340 were detonated on the open field. Altogether 91 shots were detonated. Table II shows detonation times for the shot series presented in this paper, together with some basic meteorological data.

II. METEOROLOGICAL MEASUREMENTS

A. Automatic weather stations

Aanderaa Automatic Weather Stations (AWS) were placed at three different locations, as shown in Fig. 1. To measure wind direction, and air temperature, respectively, sensor types WS2740, WD2750, and AT2775 were utilized. Wind speed and temperature sensors have accuracies of ±0.2 m/s and ±0.1 K, respectively. Threshold speed of the

TABLE I. Ground data from Haslemoen. Bottom layer (No. 4) is included as a half-space in CAPROS.

| Forest site | | | | Open field site | | |
Layer No.	Thickness (m)	Density (kg/m³)	P-wave speed (m/s)	Layer thickness (m)	Density (kg/m³)	P-wave speed (m/s)
1	0.1 (frost crust)	2100	3000	0.5 (frost crust)	2100	3000
2	0.9	1600	260	0.3	1900	1800
3	1.5	1600	320	1.7	1600	260
4	2.5	1600	420	2.5	1600	300

	On forest site		On open field site	
Shot numbers	285–291	299–305	320–326	334–340
Date	21 February	22 February	23 February	23 February
Time (UTC+1)	12:48–13:21	14:18–14:58	10:28–11:02	13:04–13:47
Charge weight	1 kg	1 kg	8 kg	8 kg
Average wind speed at 10 m	1.5 m/s	2.5 m/s	10.0 m/s	6.0 m/s
Gust values at 10 m	2.0 m/s	3.0 m/s	16.0 m/s	7.5 m/s
Air temperature at 10 m	−3 °C	0 °C	3.5 °C	2 °C

anemometers is 0.5 m/s. Sampling levels for wind speed and air temperature were 2, 10, 14, 18, and 24 m. In the open field, the same parameters were sampled at heights of 2, 5, and 10 m. Wind direction was sampled at 30 m above ground at the north forest tower, in addition to wind speed and air temperature. Meteorological data presented here are 10-min averages.

B. Sounding of the atmospheric boundary layer

A Tethersonde was used to measure atmospheric boundary layer variables. It was launched at the open field close to the open field tower (Mo) shown in Fig. 1. The Tethersonde was run almost continuously during the trial, and most of the time only the lower 500 m of the atmosphere was surveyed. Relevant Tethersonde data are wind speed and direction, air temperature, and height above ground. The latter was calculated from the relative air pressure. Sampling interval was approximately 11 s. A compromise had to be found between good height resolution, and as small time difference as possible between upper and lower measurements. Each balloon flight cycle up to 400–500 m took about 30 min, so that height resolution was better than 6 m. In a period on 23 February, while shots 306–319 were detonated, the Tethersonde could not be launched due to high wind speeds (>10 m/s). At high wind speeds, besides safety problems, a hysteresis effect will be present. The balloon drifts with the wind while ascending and against the wind while descending. This causes the wind speed measured during ascent to be too low and vice versa. This error can be minimized by decreasing balloon climbing speed. An average of ascent and descent data must be made. Since the thermometer is somewhat slow, this must be done in situations with steep vertical temperature gradients as well.

C. Turbulence measurements

An ultrasonic anemometer/thermometer, the Sonic Wind System (SWS-211/3K),[19] was applied during the trial. This instrument can measure fluctuations of wind and air temperature with a sampling frequency of 10 Hz. Thus turbulent fluxes of momentum and heat and variances can be deduced. In this context, only wind variance data will be presented to state the level of turbulence.

D. General weather conditions

A variety of weather conditions was experienced.[20] On 20 and 21 February, weather was fair and wind speeds were

less than 2.5 m/s at 10 m (Table II). Table II also shows gust values at the same height, that is, maximum wind speed over 2 s observed in the sampling interval. The difference between gust and average values is a measure of the variability of the wind. Balloon soundings indicated that a steady, slow drainage of cold air from the north dominated in the lowest 300 m of the atmosphere on 20 and 21 February. Consequently, variations in sound propagation conditions were controlled by the development of static stability in the boundary layer these days. In the afternoon of 22 February, when experiments were carried out at the open field, a warm front came from the south, and there was a general increase in temperature and wind speed. Time series of wind direction is shown in Fig. 2, with bearing directions toward the north and south towers indicated. Comparison with Table II reveals that most shot series were detonated while the wind was directed approximately along the measuring line.

E. Meteorological data

For each of the two propagation directions, the vector sound velocity profile, $V(z)$, is the sum of the speed of sound and the wind component in that direction,

$$V(z) = c_0 \sqrt{1 + T(z)/273.15} - U(z)\cos[\alpha(z) - \beta], \quad (1)$$

where c_0 is speed of sound at 0 °C, z is height above ground, $T(z)$ is air temperature, $U(z)$ is wind speed, $\alpha(z)$ is wind direction (from where the wind blows, 0°: From N, 90°:
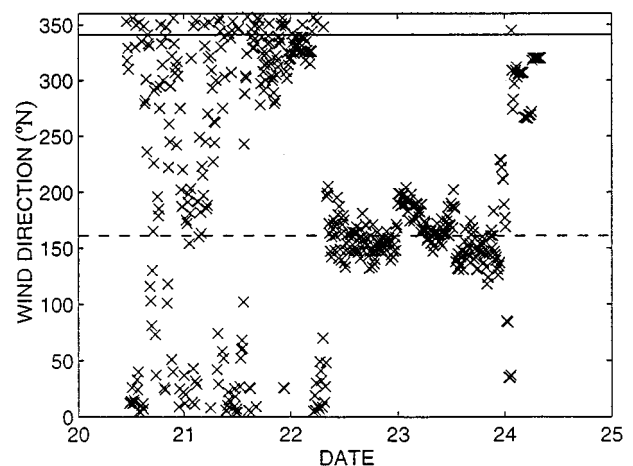
FIG. 2. Wind directions as sampled at north tower through the experiment period (crosses). Solid line: North to south tower bearing (341°). Dashed line: South to north tower bearing (161°).
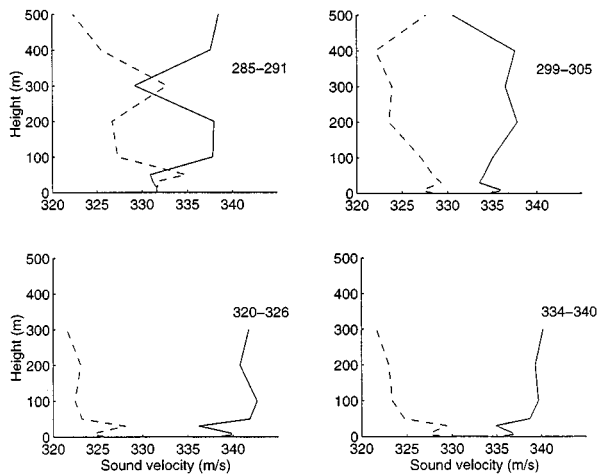
FIG. 3. Directed sound velocity profiles as resulting from combining AWS data and Tethersonde data. Solid lines: Propagation toward the north tower. Dashed lines: Propagation toward the south tower. Associated shot numbers are written on the figures.

FIG. 4. Time development of the bulk Richardson number, Ri as measured by the Tethersonde. (a) 20 February. (b) 22 February. Areas with Ri>1.0 are shaded.

From E), and β is source to receiver bearing (i.e., the angle of the source to receiver direction, relative to the north–south direction).

AWS and Tethersonde data were combined to give $V(z)$ for each shot series. Erroneous entries from the Tethersonde were corrected by interpolation or deleted; wind and temperature were smoothed over five successive entries. For specified heights 50, 100, 200, 300 m,... reports from ± 10 m were averaged and logged against observation time. For levels below 30 m, AWS data were used. Thus time interpolation could be carried out to give a best estimate of sound velocity profiles for each shot series. From this procedure, $V(z)$ is known with good height resolution in the lower 100 m of atmosphere, and with a more coarse height resolution in layers above this.

Figure 3 reveals a manifold of directed sound velocity profiles for propagation toward the north tower and south tower. By comparing with raw data (not shown), it appears that the averaging procedure described above preserves the main structure of the profiles. This structure is a decisive part of sound propagation conditions. Large vertical gradients observed in some cases are caused by wind shear, veering of wind with height, and temperature gradients. Obviously, smooth logarithmic profiles, based on extrapolation of dynamic stability conditions close to the ground, could not have been applied here. Such profiles have been used in some earlier investigations.[21] In this present investigation, it was necessary to measure directly in the relevant part of the atmosphere. Obviously, the wind speed gradients observed must generate turbulence. A straightforward measure of the level of turbulence is the turbulent kinetic energy per unit mass (TKE).[22] TKE is proportional to the sum of wind component variances, $\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$, and is defined by

$$\text{TKE} = 1/2[\overline{u'^2} + \overline{v'^2} + \overline{w'^2}]. \tag{2}$$

TKE is the fraction of the total kinetic energy which cannot be calculated from the average wind speed. Thus in a laminar flow, TKE$\equiv$0. A well-established measure of the probability of turbulence in the atmosphere as well as in other fluids is
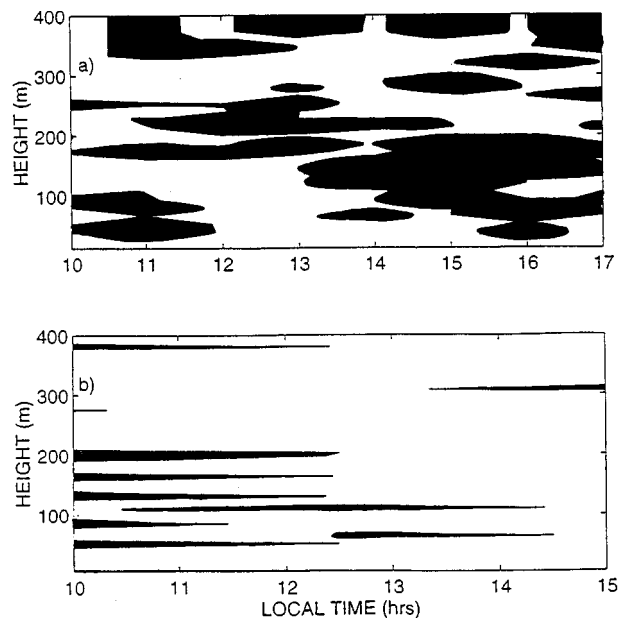
the bulk Richardson number, Ri, which characterizes dynamic stability of the flow in a layer of thickness Δz. This number is the ratio of the two most important terms in production and loss of TKE.[22] For an air layer of thickness Δz, with temperature difference ΔT and wind shear ΔU over the layer, Ri is defined by

$$\text{Ri} = -\frac{\text{buoyant production/loss of TKE}}{\text{wind shear production of TKE}} = \frac{g\Delta T\Delta z}{T_m[\Delta U]^2}, \tag{3}$$

where g is gravitational acceleration and T_m is a mean absolute temperature. Here, $T_m = 273.15$ K. It is assumed that the layer is relatively thin and that variation in humidity can be ignored, so that absolute temperature can be used instead of
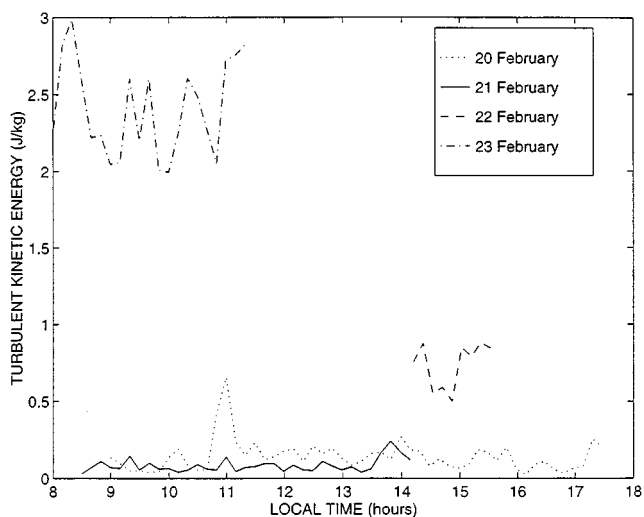
FIG. 5. Turbulence levels just above the tree tops represented by turbulent kinetic energy per unit mass (TKE) for the four experiment days. Measurements are taken by the Sonic Wind System (SWS).

TABLE III. Overview of acoustical instrumentation for each station.

	Forest trial	Open field trial
Microphones	4 Brüel & Kjær, 4147	3 B&K, 4147
Preamplifiers	4 B&K, 2639	3 B&K, 2639
Recorders	1 of 4 ch., Sony PC 204 DAT	1 of 4 ch., Sony PC 204 DAT
Analysers	2 Oni Sokki	2 Oni Sokki
Sampling levels	2, 4, 16, and 30 m	0, 2, 4, and 8 m

virtual potential temperature. It is also assumed that veering of wind across the layer is insignificant. From this definition, if the atmosphere is statically stable ($\Delta T/\Delta z > 0$), the buoyancy term will be a sink of TKE, giving a positive bulk Richardson number. As appears from (3), wind shear will always produce turbulence, which can be quantified by TKE. The range of Ri is $(-\infty, \infty)$. As a rule of thumb, if Ri < 1.0, air flow in the layer can be expected to be turbulent. If Ri > 1.0, the flow is (or becomes) laminar and, with increasing Ri, probability of turbulence decreases rapidly. Since finite differences are used, Ri can only be used as a qualitative measure of atmospheric dynamic stability. Images of Ri for 20 and 22 February are shown in Fig. 4, by the shaded areas for Ri > 1.0. Laminar flow can be expected in a large fraction of the boundary layer on 20 February, primarily in the afternoon when static stability increases in the lowest 200 m. The picture was similar on 21 February (not shown). On 22 February, a fraction of the lowest 200 m was dynamically stable in the morning, but as wind speed increases in afternoon, just a few measurements showed values of Ri > 1.0. Close to the ground, Ri is generally small due to the strong wind shear. At 23 February, values of Ri > 1.0 were not observed.

In Fig. 5, the available TKE data for each day are plotted. Measurements are taken just above the tree tops. These data confirm the general observation of more gusty wind in the last 2 days of the experiment.

III. ACOUSTICAL MEASUREMENTS

North and south towers in forest and on the open field were located as shown in Fig. 1. Table III contains a description of acoustical instrumentation for each of the four stations. Instrumentation was identical at the north and south towers. Overpressure sampling frequency was 12.8 kHz, with 4096 samples for each shot. A standard FFT procedure was used for spectral analysis.

IV. DEFINITION OF THE SOURCE

One- and 8-kg unconfined charges of C4 plastic explosives were used as sound sources. C4 has a TNT weight equivalent of 1.32.[23] Height of burst (HOB) was 2 m in all cases. A drawback of using such a relatively large explosive source is that blast propagation follows nonlinear rules in a region close to the source. Overpressure decay rate is generally much larger in the nonlinear region than in the linear

region. To estimate the overpressure decay rate from these explosions the height of burst effect must be taken into account and a scaled burst height must be applied. The HOB effect means that the apparent yield varies with height above the ground surface, and an empirical factor has to be included.[24] Based on the ANSI standard S2.20-1983,[24] Ford et al.[25] give details of these calculations. HOB factors $F_{1\,kg}$ and $F_{8\,kg}$ are 2.91 and 2.16, respectively, for detonations of 1- and 8-kg charges 2 m above ground surface. Semi-empirical formulas for positive overpressure p^+ (Pa) as a function of distance for 1- and 8-kg charges of C4 can then be shown to be

$$p^+(r)_{1\,kg} = \frac{254\,480\,6}{r^3} + \frac{465\,871}{r^2} + \frac{160\,390}{r}$$
$$\times \left[\ln\left\{ \left(\frac{r}{9.1}\right) + 3\exp\left\{-\frac{1}{3}\left(\frac{r}{81.4}\right)^{1/2}\right\}\right\}\right]^{-1/2},$$

$$p^+(r)_{8\,kg} = \frac{161\,798\,47}{r^3} + \frac{159\,886\,8}{r^2} + \frac{297\,161}{r}$$
$$\times \left[\ln\left\{ \left(\frac{r}{16.8}\right) + 3\exp\left\{-\frac{1}{3}\left(\frac{r}{150.8}\right)^{1/2}\right\}\right\}\right]^{-1/2},$$

(4)

where r is the distance in meters from the center of charge. The cubic term reflects reduction in pressure at early times due almost entirely to increasing volume, while the square term is associated with nonlinear spherical divergence, the pressure falling inversely as the surface area. The third term is a modification of the asymptotic form, the shock wave decaying toward the linear behavior of a sound wave.

A linear acoustic model cannot predict the correct overpressure decay rate in a region close to the source. The scope of this present investigation is not to study this nonlinear region, but to study, as a function of meteorology and ground interaction, sound propagation in the linear region further out. Overpressure spectrum at 20 m from the source is shown in Fig. 6 for 1-kg C4. For 8-kg charges, the central frequency is around 20 Hz (not shown). The curve is taken from Ref. 23. From Eq. (4) it can be shown that the 20-m range is well outside the nonlinear region for both charge weights (the third term is dominating), so that linear calculations are valid from this point on. Experimental data and theoretical results are presented here as transmission loss, TL (dB), defined by,

$$\mathrm{TL(dB)} = -20\log_{10}\left[\frac{\text{(total acoustic pressure at a field point)}}{\text{(acoustic pressure of direct sound field at a given reference range)}}\right].$$

(5)
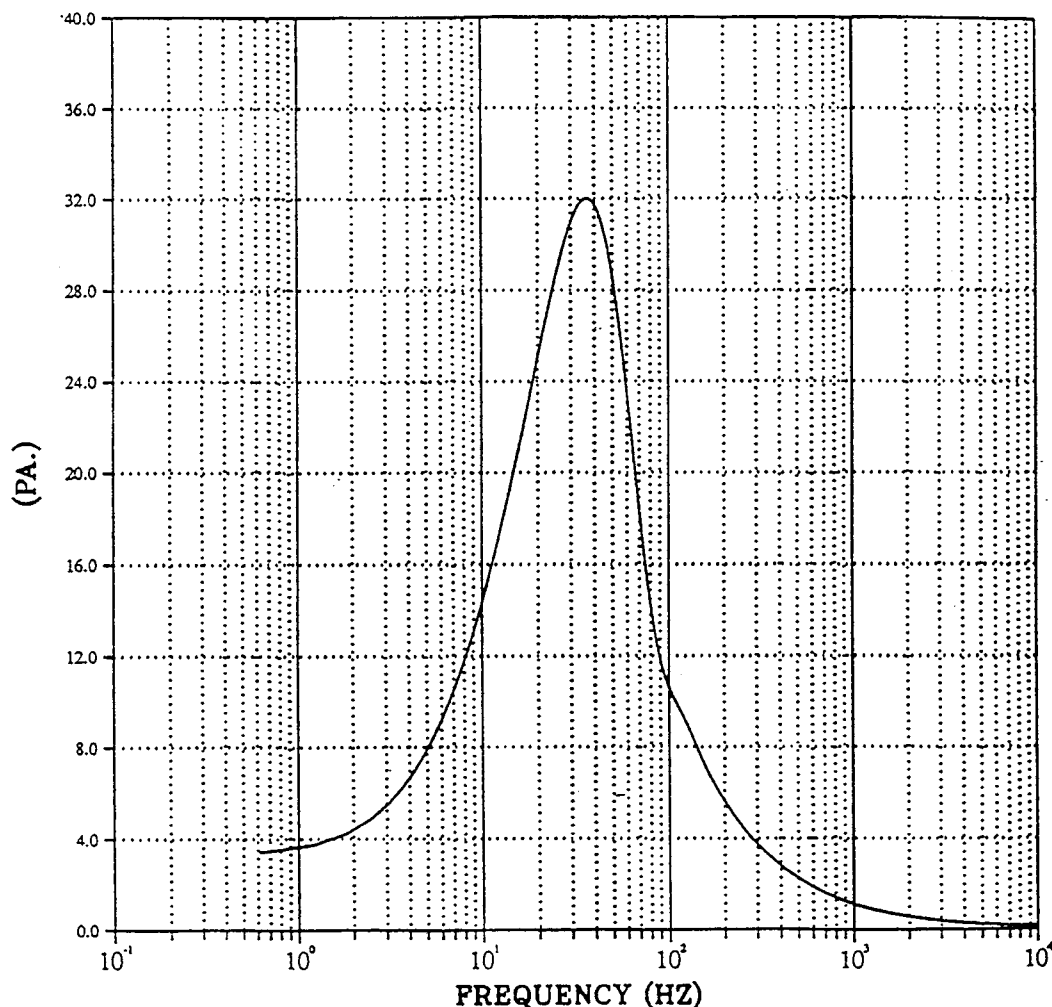
FIG. 6. Overpressure frequency content at 20 m for 1-kg C4. From Kennedy *et al.* [23]

The reference range is here set to 20 m from the source for reasons described above. Charges were detonated in series of seven from north to south at the points shown in Fig. 1. Each shot series presented here took approximately 30 min to complete, i.e., similar to the time needed for one Tethersonde flight cycle. Consequently, steady state meteorological conditions are assumed in the calculations.

V. BRIEF DESCRIPTION OF THE FAST FIELD PROGRAM CAPROS

A full wave simulation model CAPROS,[26,27] developed at the Christian Michelsen Research (CMR), has been applied for sound propagation in the atmosphere for the first time. A similar model has earlier been applied to atmospheric conditions,[16] but it has not been compared with data for low-frequency atmospheric sound propagation. However, such a comparison where experimental and theoretical waveforms are compared is now in preparation.[28] CAPROS has proven to be a powerful tool in many applications in shallow seismic and in penetration of sound fields at fluid–solid interfaces.[26,27] The CAPROS model describes linear sound propagation from a single point source (2-D propagation), or a horizontal array of point sources (3-D propagation), in a

horizontally layered, range-independent viscoelastic medium. In the present application, propagation from a single point source was studied, for which only two coordinates apply; the horizontal range, r, and height above ground, z. The model is based on a wave number integration method, using the global matrix approach. These types of solution techniques are described, e.g., in Refs. 29 and 30.

The Sommerfeld radiation condition is satisfied at infinite ranges and depths. Upper and lower layers are half-spaces. From the boundary conditions (Table IV), two and four propagation coefficients have to be found for fluid and

TABLE IV. Boundary conditions in CAPROS.

Interface	Conditions
Vacuum–Fluid	Pressure release
Vacuum–Solid	Normal and tangential stresses vanish
Fluid–Fluid	Continuity of pressure ($=-$normal stress) and vertical particle displacement
Fluid–Solid	Continuity of normal stress, tangential stress and vertical particle displacement
Solid–Solid	Continuity of normal stress, tangential stress, normal and vertical particle displacement

solid layers, respectively. The solution is represented in terms of an inverse Hankel transform. The FFP approximation is applied where only the first of the two terms in Hankel's expansion of the Bessel function is retained. It has been noted that, in some cases, both Hankel terms should be retained.[31] CAPROS allows for this alternative as well, but significant contributions from the second Hankel function, associated with the incoming wave, were not found in this investigation (not shown). The model is limited to 200 fluid layers or 100 solid layers. Relevant natural parameters for fluid layers are density, P-wave speed, P-wave absorption coefficient, and layer thickness. For solid layers, the S-wave speed and the S-wave absorption coefficient must be given in addition. CAPROS has been tested on the benchmark cases in Ref. 9, where the ground consisted of porous sand with a flow resistivity of 366 000 Pa s m^{-2}. An atmospheric layer thickness of 10 m was used, with 16 384 points in the transform integrals. Predicted TL were correct to within 1 dB out to about the 1-km range (not shown) for 10 and 100 Hz. At longer ranges, sound levels were estimated too high probably due to the assumption of a viscoelastic, nonporous ground. However, with the acoustically hard ground present in this experiment, and at these low frequencies, we believe that the assumption of a viscoelastic ground is adequate. Proper convergence testing with respect to the wave number interval and sampling rate to chose, has to be carried for each case studied.

VI. COMPARISON OF EXPERIMENTAL RESULTS AND MODEL PREDICTIONS

A. Employing meteorological data in CAPROS

When applying CAPROS, sound velocity profiles as presented in Fig. 3 were transformed into step functions. Each step represents a theoretical homogeneous, isotropic layer. Test runs showed that the atmosphere above 300 m did not affect TL along the ground, so that data from above this height were not included in later runs. Layer thicknesses were 2, 3, 5, 10, 10, 20, 50, 100, and 100 m, respectively from ground and toward increasing altitude. An air density of 1.2 kg/m^3 and an atmospheric absorption coefficient of 0.5 dB/km was applied. The real absorption coefficient is of size order 0.1–0.5 dB/km in the frequency range and climate studied here,[14] but a conservative estimate of the absorbtion coefficient ensures a numerically stable solution. The influence of absorbtion is anyway negligible compared to the other effects studied. In the viscoelastic ground, an attenuation coefficient of 20 dB/km was employed. (Theoretical and observed transmission losses are compared in Figs. 9–12.) In most cases, 8192 points was a sufficient resolution to obtain stable solutions from the wave number integration, but in some cases with steep sound velocity gradients as much as 32 768 points were required. This is indicated in figure labels. In all cases, receiver height is 4 m, and source height (HOB) is 2 m.

B. Acoustical observations

An example of observed overpressure versus time signatures for shot 305 (on the open field) is shown in Fig. 7.

The range is 100 m. Charge weight was 1 kg. As in many other observations (also in the forest), the overpressure was highest at the lowest measuring level, but there is no dramatic difference. In general, peak overpressure was slightly lower in the forest compared to open field observations, but a closer analysis suggests that it is only the highest frequency studied which shows a general tendency of excess attenuation compared to model predictions. This observation concurs with conclusions from the summer trial at Haslemoen in June 1994.[32] Figure 8 reveals that, for all ranges, the sound is most attenuated at 63 Hz, but that there is a significant standard deviation for all frequencies. This figure represents data for all shots. Standard deviation of the TL is also by far largest at 63 Hz, indicating that this frequency is most affected by weather variations. If observations from forest and open field are separated, the tendency is similar (not shown). Dashed lines are inverse square spreading loss curves. The average data do not seem to show any evidence of the cylindrical spreading loss anomaly reported in Refs. 12 and 13 for strongly refracted downwind infrasonic propagation.

C. Cases with low wind speed

Theoretical and experimental results from some interesting situations are presented here. The criterion has not been that the best fits were to be presented. Some situations with steep wind speed gradients are chosen, so that effects of meteorology will be most distinct. Profiles are characterized by succeeding layers with gradients of opposite signs. This is mostly caused by veering of the wind. Of course, for such a complicated profile, it cannot be predicted intuitively how sound level along the ground will be affected. As Fig. 9 reveals, for shots 285–291 (in the forest), deviation between measured and predicted values is less than 10 dB for the three lowest frequencies, but at 63 Hz, a deviation of up to 45 dB is observed at the longest ranges. Weather conditions were stationary with low wind speed (1.5 m/s) and the turbulence level was very low as shown in Figs. 4 and 5. At 63 Hz the excess attenuation is large in both directions. This suggests that the deviation is caused by the presence of forest, not meteorology. This is a common observation for all shots in the forest.

Next results presented are observations at the south tower for shots 299–305 (on the open field), shown in Fig. 10. Both data and predictions lie on smooth lines, and the deviation is less than 4–5 dB in most cases, including the highest frequency, even though a large fraction of the atmosphere was dynamically unstable this day (Figs. 4 and 5).

D. Cases with higher wind speed

Another interesting result is shown in Fig. 11. The fit is not satisfactory for many shots, but both experimental data and particularly the predictions show great ''dips.'' Depth and sharpness of the ''dips'' seems to increase with frequency. CAPROS indicate the presence of shadow zones in this case, but the location is not predicted accurately. These shots were detonated while the average wind speed at 10 m was between 10 m/s, with gust values up to 16 m/s. This high wind speed also caused very steep velocity gradients in
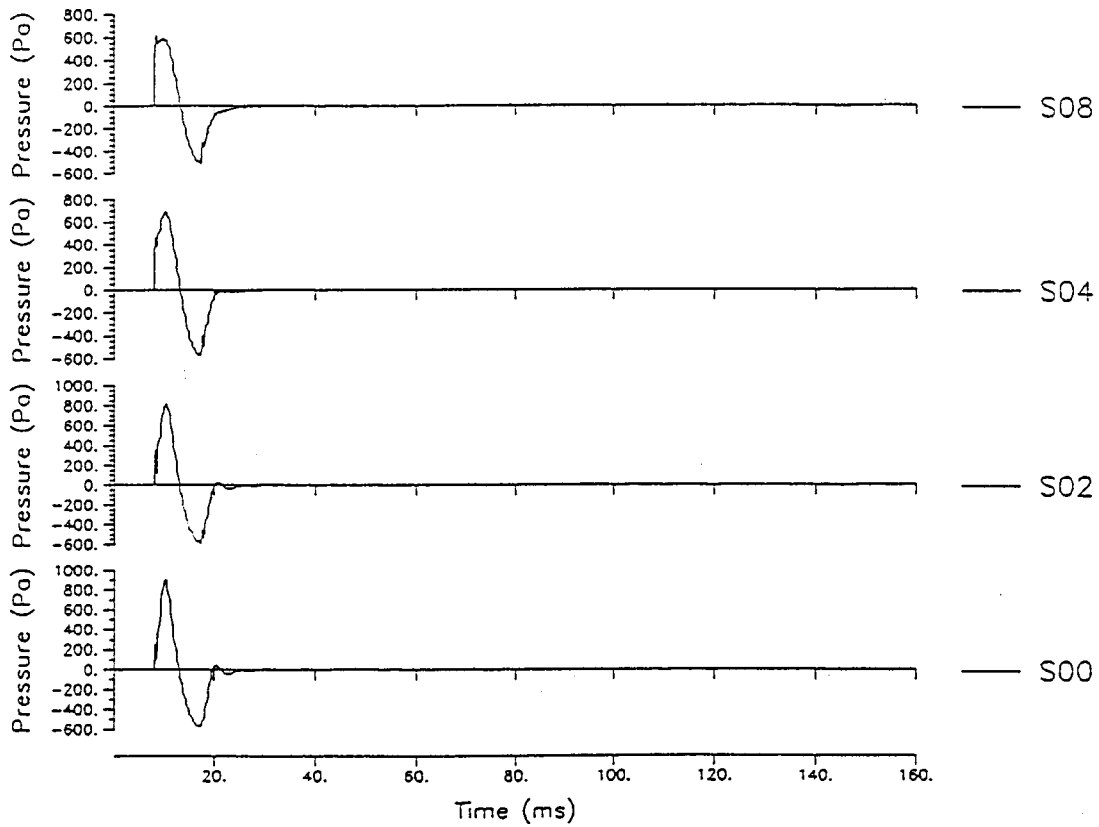
FIG. 7. Overpressure versus time at 0, 2, 4, and 8 m for shot 305 measured at south tower (100-m range). Measuring heights are shown on the right.
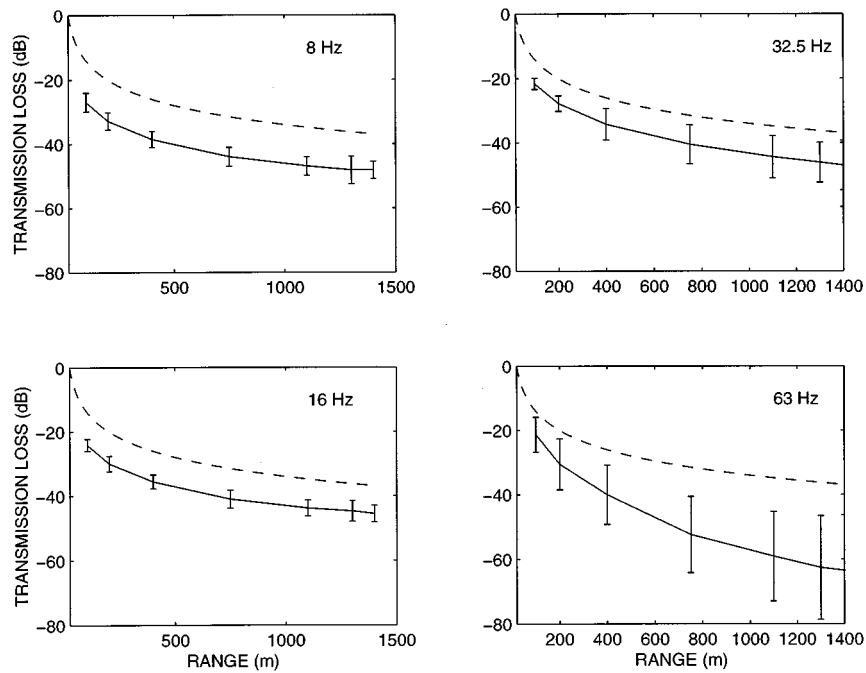
FIG. 8. Average transmission loss relative to 20 m versus range at four frequencies for all shots during the trial; both south and north tower measurements are included. Bars indicate the standard deviation at each range. Dashed lines are inverse square spreading loss curves. Source height is 2 m, while receiver height is 4 m.
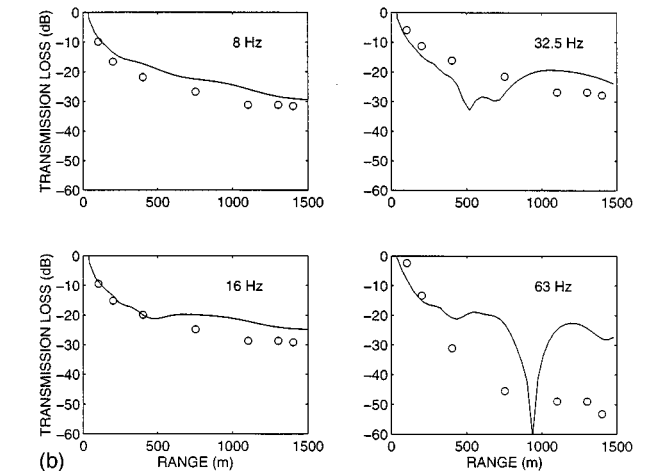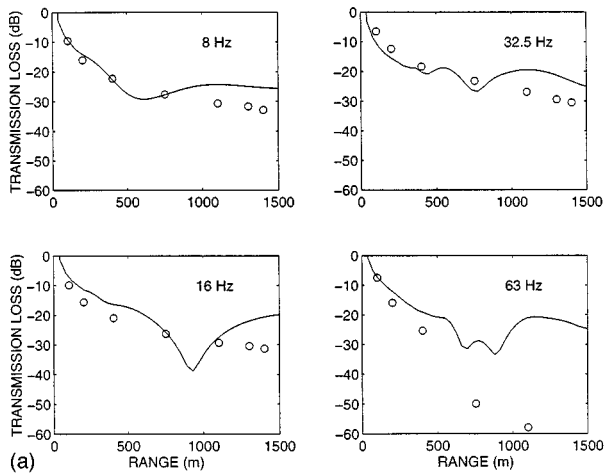
FIG. 9. Predicted (line) and measured (circles) transmission loss versus range, shots 285–291; (a) at north tower; (b) at south tower. 8192 integration points.

the lowest part of the atmosphere. We suggest that the deviations observed are mostly caused by temporal variability of the wind.[33] It is obvious that the assumption of steady state atmospheric conditions is not valid here.

Finally, data from shots 334–340 are shown in Fig. 12
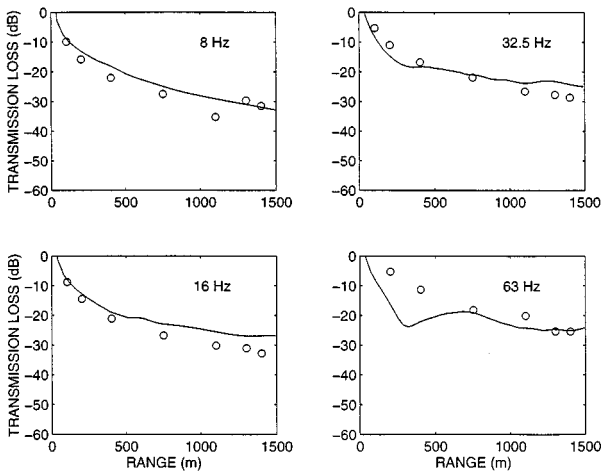
FIG. 10. (a) Predicted (line) and measured (circles) transmission loss versus range, shots 299–305, at south tower. 8192 integration points.
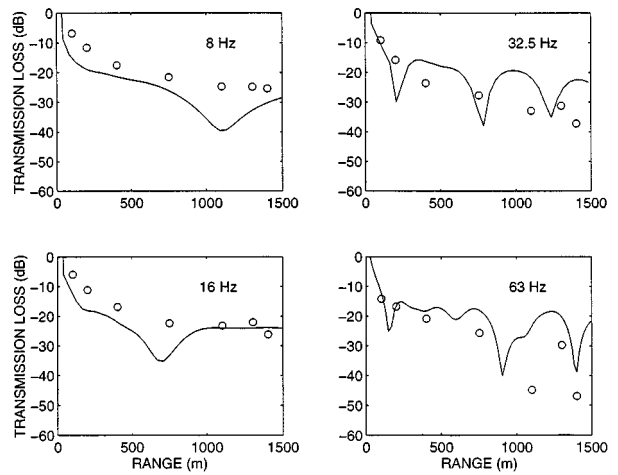
FIG. 11. Predicted (line) and measured (circles) transmission loss versus range, shots 320–326, at north tower. 32 768 integration points.

as observed at the south tower. V is mostly decreasing with height in this case (Fig. 3), except in the lowest 30 m. Average wind speed at 10 m is now reduced to about 6 m/s, and gust values are only about 1–2 m/s larger than average values. Even if gradients of V are still steep, curves are more smooth, and deviations are acceptable in most cases. It seems like predictions are satisfactory when variability of the wind is not too high.

VII. CONCLUDING REMARKS

A fast field program, CAPROS, has been used to predict propagation of sound at 8, 16, 32.5, and 63 Hz. These predictions were compared with experimental data at ranges from 100 to 1400 m. Sound sources were 1- and 8-kg charges of C4 explosives. A viscoelastic ground was assumed. This is considered adequate, since the ground was frozen and since only low frequencies were studied. Predictions were based on average directed sound velocity profiles that were calculated from *in situ* meteorological measurements in the atmospheric boundary layer. The resulting pro-
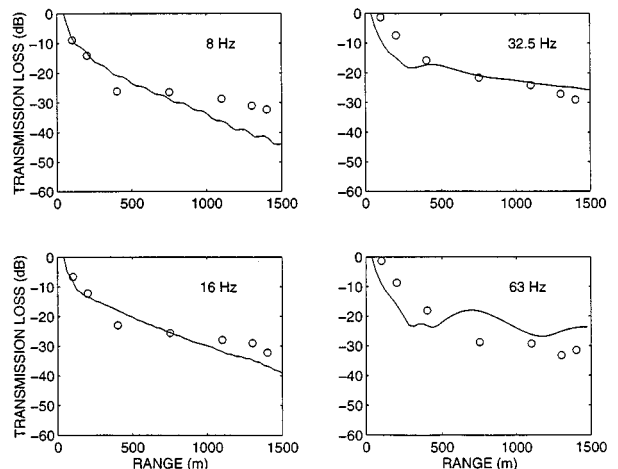
FIG. 12. Predicted (line) and measured (circles) transmission loss versus range, shots 334–340, at south tower. 8192 integration points.

files were complex and represented a challenge for CAPROS that it seems to have managed. Test runs of CAPROS showed that meteorological data to a height of 300 m had to be included to give stable predictions of transmission losses along ground. Predictions based on average profiles seems meaningful in this frequency range, even if the turbulence level is relatively high. Only in one episode with very high variability of the wind, the deviation from model to experimental data is not acceptable. Most of the data presented are from an open field. In addition, it is shown that a pine forest probably causes an excess attenuation at 63 Hz that is not observed at the lower frequencies. In this case both source and receivers were placed in the forest. The standard deviation of transmission losses along ground increases with frequency and propagation range, but all frequencies seems to be affected by meteorological variations. In general, there seems to be more ''structure'' on theoretical curves than on experimental data. With few exceptions, experimental data points reveal smooth attenuation with distance. The ''dips'' observed on theoretical predictions are generally not confirmed by observations. These ''dips'' probably result from the stratified nature of the model, while real world profiles are less coarse.

Meteorological variability was a source of error during part of the trial. Transient changes in atmospheric profiles might have been present during each balloon ascent–descent, which is a probable reason for some of the deviations observed.[20] Therefore, other measurements have been carried out under extremely stationary weather conditions, so that this error is minimized. A presentation of these data together with a comparison with FFP and parabolic equation predictions is in preparation. Also, a poro-elastic ground model will be applied to get a realistic prediction of air–ground interaction in cases where a highly absorptive surface such as a thick snow layer is present.

Information about the experimental sound propagation series *Norwegian Trials* is available in Refs. 34–36.

ACKNOWLEDGMENTS

[1] K. B. Rasmussen, ''The influence of wind on sound propagation over porous ground,'' International Congress on Acoustics (ICA'95), Trondheim, Norway, 255–258 (1995).

[2] K. B. Rasmussen, ''Sound propagation over ground under the influence of a sound speed profile in the atmosphere,'' J. Sound Vib. **139**, 71–81 (1990).

[3] A. L'Espérance, Y. Gabillet, and G. A. Daigle, ''Outdoor sound propagation in the presence of atmospheric turbulence: Experiments and theoretical analysis with the fast field program algorithm,'' J. Acoust. Soc. Am. **98**, 570–579 (1995), and references herein.

[4] R. Raspet and R. K. Wolf, ''Application of the fast field program to the prediction of average noise levels around sources,'' Appl. Acoust. **27**, 217–224 (1989).

[5] D. K. Wilson, ''Sound field calculations in a stratified, moving medium,'' J. Acoust. Soc. Am. **94**, 400–407 (1993).

[6] Y. L. Li, S. J. Franke, and C. H. Liu, ''Numerical implementation of an adaptive fast-field program for sound propagation in layered media using the chirp z transform,'' J. Acoust. Soc. Am. **89**, 2068–2075 (1991).

[7] Y. L. Li, M. J. White, and S. J. Franke, ''New fast field programs for anisotropic sound propagation through an atmosphere with an wind velocity profile,'' J. Acoust. Soc. Am. **95**, 718–726 (1994).

[8] M. West, R. A. Sack, and F. Walkden, ''The fast field program (FFP). A second tutorial: Application to long range sound propagation in the atmosphere,'' Appl. Acoust. **33**, 199–228 (1991), and references herein.

[9] K. Attenborough, S. Taherzadeh, H. E. Bass, R. Raspet, G. R. Becker, A. Güdesen, A. Chrestman, G. A. Daigle, A. L. Espérance, Y. Gabillet, K. E. Gilbert, Y. L. Li, M. J. White, P. Naz, J. M. Noble, and H. A. J. M. van Hof, ''Benchmark Cases for outdoor sound propagation models,'' J. Acoust. Soc. Am. **97**, 173–191 (1995), and references herein.

[10] J. S. Robertson, W. L. Siegman, and M. J. Jacobson, ''Low-frequency sound propagation modeling over a locally reacting boundary with the parabolic approximation,'' J. Acoust. Soc. Am. **98**, 1130–1137 (1995).

[11] R. N. Tedrick and R. C Polly, ''Measured acoustic propagation parameters in the Mississippi Test Operations Area,'' NASA TM X-1132, Aug. 1965.

[12] H. H. Hubbard, K. P. Sheperd, and W. L. Willshire, ''Results of simultaneous acoustic measurements around a large horizontal axis wind turbine,'' J. Acoust. Soc. Am. **80**, 528 (1995).

[13] W. E. Zuromski and W. L. Willshire, ''The acoustic field of a point source in a uniform boundary layer over an impedance plane,'' Paper AIAA-86-1923, presented at AIAA Aeroacoustic conference, July 1986.

[14] C. Larsson, ''Swedish climate for absorption of sound by the atmosphere,'' Internoise'96, 601–604, 1996.

[15] C. Larsson, ''Meteorological effects on sound propagation,'' Current Topics in Acoust. Res. **1**, 91–99 (1994), Council of Scientific Information, India.

[16] A. Güdesen, ''Application of the SAFARI model to sound propagation in the atmosphere,'' J. Acoust. Soc. Am. **87**, 1968–1974 (1990).

[17] A. M. Kaynia and C. Madshus, ''Blast propagation through forest—NOR95/1 tests,'' NGI report 515137-4, Norwegian Geotechnical Institute (1995).

[18] K. H. Stokoe, G. R. Rix, and S. Nazarian, ''*In situ* seismic testing surface waves,'' 12th International Conference on Soil Mechanics and Foundation Engineering, Rio de Janeiro, **1**, 331–334, 13–18 August 1989.

[19] L. R. Hole, ''An experimental study of validity of two methods for estimating sensible heat flux following a step change in surface temperature,'' Contrib. Atmos. Phys. **69**, 409–415 (1996).

[20] L. R. Hole, Y. T. Gjessing, T. de Lange, and J. W. Reed, ''Meteorological measurements and conditions during *Norwegian Trials*,'' invited paper, Noise Control Eng. J. (December 1996).

[21] A. L'Espérance, J. Nicolas, D. K. Wilson, D. W. Thomson, Y. Gabillet, and G. Daigle, ''Sound propagation in the atmospheric surface layer: Comparison of experiment with FFP predictions,'' Appl. Acoust. **40**, 325–346 (1993).

[22] R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Kluwer Academic, Dordrecht, The Netherlands, 1988), pp. 175–180.

[23] L. Kennedy, S. Hikida, and Robert Ekler, ''Overpressure and dynamic pressure waveforms for small C4 charge detonations,'' Report No. SSS-DFR-94-14507 Maxwell S-Cubed Division, Albuquerque, New Mexico (1994). Report performed for NDCS.

[24] ''American National Standard for estimating airblast characteristics for single point explosions in air,'' ANSI S2.20-1983, Acoustical Society of America, New York (1983).

[25] R. Ford, D. J. Saunders, and G. Kerry, ''The acoustic pressure waveform from small unconfined charges of plastic explosive,'' J. Acoust. Soc. Am. **94**, 408–417 (1993).

[26] P. Lunde, P. Meldal, and M. Vestrheim, ''Full-wave elastic modelling for seismic detection of shallow gas,'' paper presented at the 49th Meeting

and Technical Exhibition for the European Association of Exploration Geophysics (EAEG'87), Belgrade June 9–12, 1987. CMI Ref. No. 1987:67, Christian Michelsen Research, Bergen (1987).

[27] P. Lunde, ''Interference effects in sound fields transmitted at fluid–fluid and fluid–solid Interfaces,'' paper presented at Scandinavian Cooperation Meeting in Acoustics/Hydrodynamics, Ustaoset, Norway, 27–30 January 1991. CMI ref. CMI-91-A10001. Christian Michelsen Research, Bergen (April 1991).

[28] L. R. Hole, ''An experimental and theoretical study of propagation of acoustic pulses in a strongly refracting atmosphere,'' Appl. Acoust. (to be published).

[29] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Scmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994).

[30] H. Schmidt and F. B. Jensen, ''Efficient numerical solution technique for wave propagation in horizontally stratified environments,'' Comput. Math. Appl. **11**, 699–715 (1985).

[31] Y. L. Li and M. J. White, ''A note on using the fast field program,'' J. Acoust. Soc. Am. **95**, 3100–3102 (1994).

[32] R. D. Ford, ''Report on the measurement of blast noise propagation in the forest, Haslemoen, Norway, 9–16 June 1994,'' Rep. No. Sal/FT/32, University of Salford, UK (1995).

[33] J. S. Robertson, ''Numerical simulation of intensity fluctuations caused by wind gusts,'' J. Acoust. Soc. Am. **87**, 1353–1355 (1990).

[34] G. Kerry, ''An overview of the long range impulse sound propagation measurements made in Norway,'' Internoise'96, 583–588 (1996), and references herein.

[35] R. L. Guice and A. Jenssen, ''Impulse noise measurements in a forest during summer and winter conditions—An overview of measurements,'' Internoise'96, 2589–2594 (1996).

[36] L. R. Hole and R. L. Guice, ''Norwegian Trials, the experiment and the database,'' paper to be presented at the 5th Congress on Sound and Vibration, December 1997, and references therein.

The use of an array to explain the sound characteristics of secondary small plate tones produced by the impingement of an axisymmetric choked jet

Brenda Henderson[a)] and Alan Powell

University of Houston, Department of Mechanical Engineering, Houston, Texas 77204-4792

The secondary small plate tones, referred to as S2 tones by Henderson and Powell [J. Sound Vib. **168**, 307–326 (1993)], occur over a more limited range of pressures and nozzle-to-plate spacings than do the primary small plate tones [see B. Henderson and A. Powell, J. Acoust. Soc. Am. **99**, 153–162 (1996)] but are still quite intense. Secondary tones are usually produced when plates are positioned near the location of the second cell ending in the free jet. Their power spectra usually contain a single discrete peak, the fundamental frequency, although occasionally the first harmonic is also present. Timed shadowgraph photographs indicate that the axial location and shape of the shock waves change randomly with time and are not connected with the production of tones. The motion of the shock wave is small in comparison to the motion observed for the primary small plate tones. Results of experiments conducted with a nozzle reflector indicate that these tones are part of a feedback loop to the nozzle. A four-element simple-source array with the array elements located at the approximate shock wave positions in the jet and the flow along the plate has been found to explain the measured frequencies and directivities. © *1997 Acoustical Society of America.* [S0001-4966(97)02808-7]

PACS numbers: 43.28.Ra, 43.50.Nm [LCS]

INTRODUCTION

Three types of tones are produced by the impingement of a choked jet on a normal flat plate [see Powell and Henderson (1990) and Henderson and Powell (1993)]. All three types of tones often have sound-pressure levels that exceed the local broadband noise by 20 dB. Two types of tones are associated with impingement on small plates where small plates are defined as plates with diameters less than or equal to two nozzle diameters and one type of tone is associated with impingement on large plates. Large plate tones can occur for plate diameters greater than or equal to two nozzle diameters and have different frequency characteristics than both types of tones associated with small plates. Small plate tones can occur for plate diameters less than or equal to two plate diameters. Although primary and secondary small plate tones are associated with the same diameter plates, the sound-production mechanisms for these two types of tones are quite different [see Henderson and Powell (1996)] and there are observed differences in the frequency characteristics associated with each type of small plate tone. A more detailed description of the frequency characteristics associated with each type of tone may be found in Henderson and Powell (1993), where the large plate tones were referred to as L1, L2, and L3 tones, the primary small plate tones were referred to as S1 tones, and the secondary small plate tones were referred to as S2 tones. The current investigations are concerned with gaining information about the sound-production mechanism associated with the secondary small plate tones, S2 tones. These tones are referred to as secondary tones as they occur over a more limited range of operating conditions than the primary small plates tones; however, they can be as much as 30 dB above the local broadband noise.

The discrete tones associated with the free choked jet, known as choked jet screech, are produced by the interaction of sinuous stream disturbances (instabilities) with the regularly spaced shock waves in the choked jet [see Powell (1953)]. Once produced, the sound waves travel back to the nozzle exit where they interact with the flow to create stream disturbances. This completes the feedback loop. Powell (1962) explained the frequencies of these tones by calculating the wavelength which resulted in maximum pressure reinforcement at the nozzle exit from three and from four simple sources located in the same positions as the regularly spaced shock waves. Harper-Bourne and Fisher (1974) also used a simple source array to effectively describe the sound produced by broadband shock associated noise produced by a free jet operating above the critical pressure ratio. This type of sound is produced by the interaction of turbulence with the regularly spaced shock waves in the jet.

The production of discrete tones by impinging jets was first reported by Marsh (1961). Ho and Nosseir (1981) hypothesized that the tones produced by a subsonic jet impinging on a flat plate were part of a feedback loop to the nozzle and were the result of the interaction of large coherent structures (vortices) with the plate. Nosseir and Ho (1982) inferred that most of the sound was produced by the interaction of the coherent structures with the plate, and Didden and Ho (1985) observed large ring vortices in the subsonic flow using smoke visualization techniques.

Powell (1988) investigated the supersonic impinging jet at ''moderate'' expansion ratios and hypothesized that the

---

tones produced by impingement on large plates were part of a feedback loop to the nozzle, while tones produced by impingement on small plates were related to the high harmonics of the Hartmann whistle and were not part of a feedback loop to the nozzle. A Hartmann whistle consists of a supersonic jet exhausting into a cylindrical cavity [see Hartmann and Trolle (1927)]. Large oscillations of the standoff shock wave which occurs in front of the plate were noted by Powell when small plate tones were produced. These oscillations appeared to be similar to the oscillations of the standoff shock wave in the Hartmann whistle. In the investigations of Henderson and Powell (1993), two types of small plate tones were identified and both appeared to be part of a feedback loop to the nozzle. These tones were referred to as S1 and S2 tones (primary and secondary small plate tones) and had different frequency characteristics.

In the investigations into the production mechanism associated with primary small plate impinging tones (S1 tones) conducted by Henderson and Powell (1996), it was found that S1 tones appeared to be generated by contractions of the jet boundary caused by the large oscillations of the first shock wave and its subsequent interaction with the standoff shock wave (the shock wave in front of the plate). Primary small plate tones did not appear to be part of a feedback loop to the nozzle and there were significant differences between the flows producing the Hartmann whistle tones and the primary small plate impinging tones.

The references mentioned here are limited. A more detailed discussion of previous investigations into the tones produced by impinging jets may be found in Powell (1988) and Henderson and Powell (1993, 1996).

The present investigation is concerned with the sound-production mechanism of the secondary small plate tones referred to as S5 tones by Powell and Henderson (1990) and as S2 tones by Henderson and Powell (1993, 1996). A brief description of the apparatus used in the experimental investigations may be found in Sec. I. A spectral analysis of the secondary small plate tones will be given in Sec. II. In Sec. III, flow visualization results will be presented. Investigations conducted with a convergent–divergent nozzle in place of the convergent nozzle will be discussed in Sec. IV. Results of reflector tests used to determine the possibility of feedback to the nozzle will be given in Sec. V. In Sec. VI, a hypothesized sound-production model for the secondary small plate tones will be developed. A discussion of the results will be given in Sec. VII, and conclusions will be presented in Sec. VIII.

I. APPARATUS

A schematic of the apparatus is shown in Fig. 1. Air exhausting from a convergent nozzle with an exit diameter equal to 0.80 cm impinges on a flat plate supported on a traversing mechanism. The nozzle and traversing mechanism are enclosed in a small anechoic chamber with removable walls which allow light rays from a schlieren and shadowgraph system to pass through the jet. The knife edge in the schlieren system is oriented vertically. Frequency measurements are made in the far field with a microphone supported on the same traverse and centered below the flat plate. The
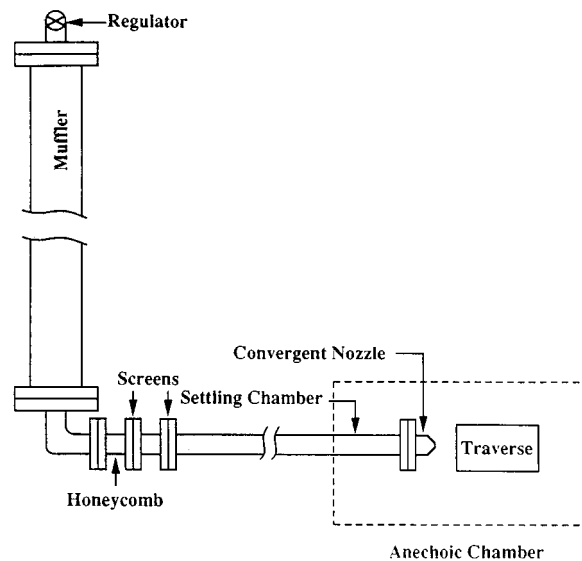
FIG. 1. A schematic of the experimental apparatus.

frequency signals are analyzed on an FFT analyzer. A detailed description of the apparatus may be found in Henderson and Powell (1996).

II. SPECTRAL ANALYSIS

A plot of the normalized acoustic wavelength (λ/d) versus the normalized nozzle-to-plate spacing (h/d) is shown in Fig. 2 for a stagnation pressure ratio of $p_0/p_a = 3.72$, where p_0 is the stagnation pressure, p_a is the ambient pressure, h is the nozzle-to-plate spacing, λ is the acoustic wavelength, and d is the nozzle exit diameter. This plot includes only fundamental frequencies produced by impingement on circular plates with diameters equal to $D/d = 1$, $D/d = 2$, and $D/d = 3$, where D is the plate diameter. The lines labeled L are associated with large plate tones and the lines labeled S1 and S2 are associated with small plates tones. When the plate diameter is greater than or equal to 4 nozzle diameters, only large plate tones occur and fall along the same lines as indicated by L1 and L2 in the figure. For plate diameters equal to $D/d = 2$, both small plate tones and large plate tones may
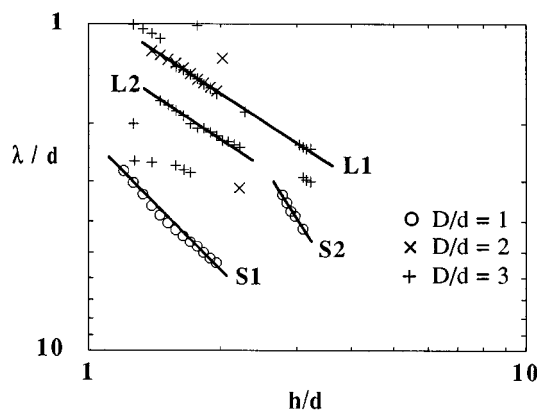
FIG. 2. Wavelength data for a jet with $p_0/p_a = 3.72$ impinging on plates with sizes equal to $D/d = 1$, $D/d = 2$, and $D/d = 3$.

B. Henderson and A. Powell: Array model for secondary tones
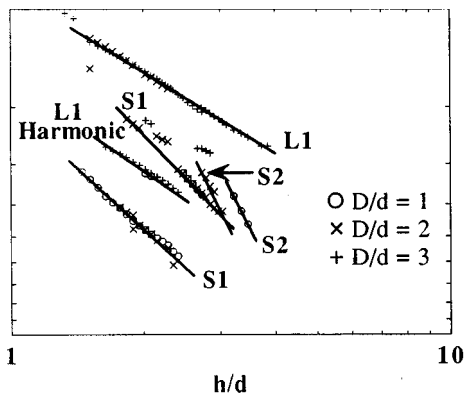
FIG. 3. Wavelength data for a jet with $p_0/p_a = 4.40$ impinging on plates with sizes equal to $D/d = 1$, $D/d = 2$, and $D/d = 3$.

occur. The data falling along the S1 lines are the primary small plate tones discussed by Henderson and Powell (1996) and the data falling along the S2 line are the secondary small plate tones. This pressure ratio ($p_0/p_a = 3.72$) is the lowest pressure ratio at which S2 tones are produced. Generally, S1 tones occur at lower pressures than do S2 tones.

The wavelength data presented in Fig. 3 were obtained at a stagnation pressure ratio of $p_0/p_a = 4.40$. By comparison with the previous plot, it may be seen that a second S2 tone occurs with an increase in pressure. When the stagnation pressure increases to $p_0/p_a = 4.74$, all small plate tones are S1 tones although some of the tones associated with the $D/d = 2$ plate appear to fall along somewhat steeper lines as expected for the S2 tones.

The frequency plots indicate that the S2 tones usually occur when the plate is located near where the second cell ending would occur in the free jet. A few of the tones produced when the plate is located near the ending of the first cell in the free jet may be S2 tones as well. This is an indication that the production of these tones is sensitive to the structure of the impinging jet since the structure of the impinging jet, although different, is related to the structure of the free jet. Henderson and Powell (1996) found that the production of S1 tones is also sensitive to the structure of the jet, but these tones are usually associated with a different range of plate spacing than the S2 tones.

A typical power spectrum associated with the S2 tones is shown in Fig. 4. As is usual for these tones, only a single discrete peak occurs although there are occasions where a
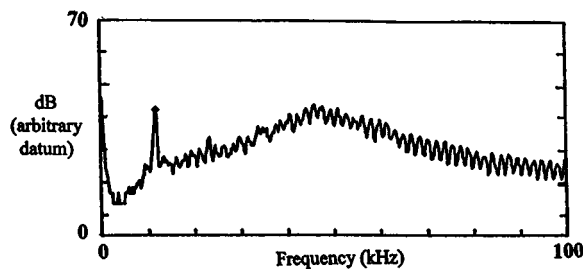
FIG. 4. A typical power spectrum for the secondary (S2) tones. The data for this plot were taken at $p_0/p_a = 4.06$, $D/d = 1$, and $h/d = 3.10$ ($\lambda/d = 3.8$).

second peak, believed to be the first harmonic, is also present. This is in contrast to the S1 tones which usually are associated with many significant harmonics in the power spectra [see Fig. 3 of Henderson and Powell (1996)]. It may be seen from Fig. 4 that the S2 tones have quite intense peak sound-pressure levels which exceed the local broadband noise by approximately 20 dB.

III. FLOW VISUALIZATION

Timed shadowgraph photographs were taken with the timing circuit described by Henderson and Powell (1996). The light source in the optical system was triggered by a signal from a microphone placed in the far field. The microphone signal was displayed on an oscilloscope and after choosing different time delays on the displayed wave form, a signal was sent to the light source through a gated output on the oscilloscope.

The results of the timed shadowgraph experiments for an operating condition where only S2 tones occur are shown in Fig. 5. The photographs in this figure represent four photographs in a set of 11 photographs taken at equal time intervals during one cycle of the fundamental oscillation (but not the same cycle).

In Fig. 5, the first shock wave in the jet has the same shape and is in the same downstream location as the first shock wave in the free jet at the same stagnation pressure. This first shock wave remains quite stationary and has approximately the same shape in all of the photographs. The second shock wave is also present in the free jet but its position and shape have been altered by the presence of the plate. The downstream location and dome shape of the second shock wave change somewhat randomly in these photographs. The variation in the dome shape of a second shock wave was noted by Powell (1988). The third shock wave, a standoff shock wave, disappears when the second shock wave moves slightly downstream; otherwise its position and shape remain almost unchanged throughout the cycle. The shape of the standoff shock wave is most likely annular about a subsonic central point, although it is not possible to determine this from these photographs.

Henderson and Powell (1996) found that, when S1 tones are produced, the standoff shock wave becomes dome shaped during a portion of the cycle and the first Mach disk accelerates downstream until it interacts with the standoff shock wave. Only two shock waves occur in the jet upstream of the plate and the Mach disk of the first shock wave has a much larger radius than the Mack disk of the first shock wave occurring when S2 tones are produced. In all of the photographs taken where S2 tones are produced, the standoff shock wave is never dome shaped and does not appear to interact with the upstream shock wave as it does when S1 tones are produced. The jet structures associated with the S1 and S2 tones are, therefore, significantly different.

In the shadowgraph photographs shown in Fig. 5, and the schlieren photograph shown in Fig. 6, there appear to be stationary shock waves in the outer regions of the deflected flow just downstream of the plate face. In all of the photographs in Fig. 5, there are two of these shock waves in nearly

(a)

(b)

(c)

(d)
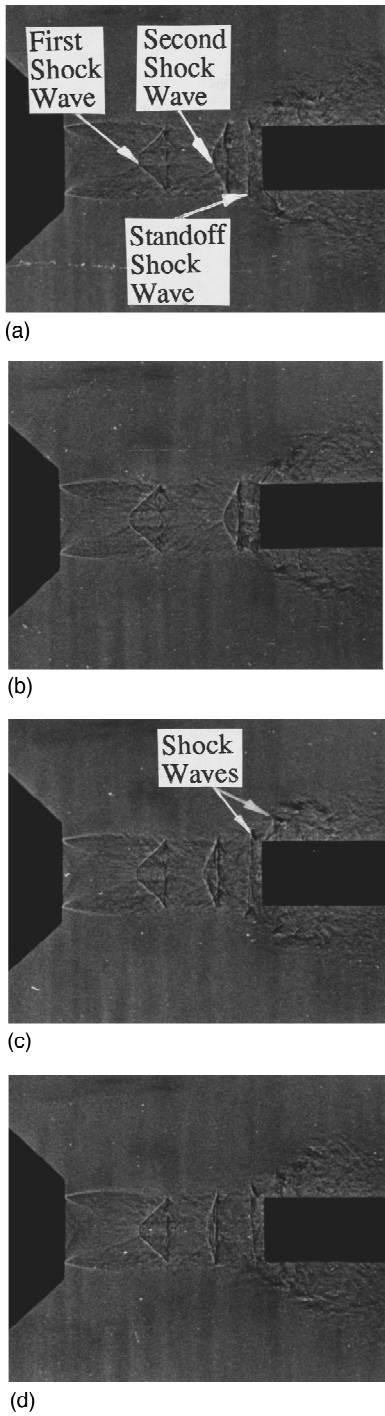
FIG. 5. Timed shadowgraph photographs taken at $p_0/p_a = 4.14$, $h/d = 3.04$, and (a) $t = 0$, (b) $t = 30$ μs, (c) $t = 60$ μs, and (d) $t = 90$ μs. The oscillation period is 100 μs.

the same fixed position. In Sec. VI, these shock waves will be shown to be important to the sound-production mechanism of the S2 tones.

The radiated wave fronts in the schlieren photograph in Fig. 6 are not as narrow as the wave fronts associated with S1 tones and, therefore, are expected to be less impulsive in nature. This is consistent with the lack of harmonics in the power spectra associated with these tones. More importantly, the curvatures of the wavefronts are not all the same, indicating that there is more than one sound source in the flow. If
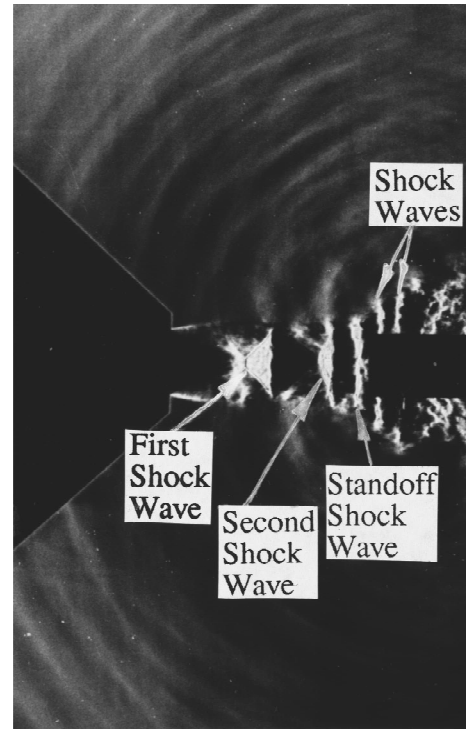
FIG. 6. Schlieren photograph taken at $p_0/p_a = 4.06$ and $h/d = 3.04$.

normals are drawn to the wavefronts, it appears that the sound originates from the outer edge of the second shock wave (the standoff shock wave) and from the stationary shock waves in the flow along the plate.

A plot of the axial motion of the shock waves over one cycle (but not the same cycle) is shown in Fig. 7. As mentioned above, the first shock wave is almost stationary
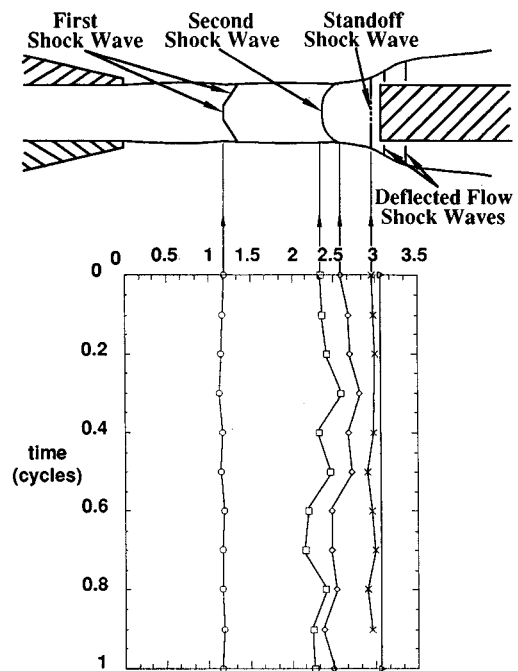
FIG. 7. Motion of the first, second, and standoff shock waves over one cycle. The data are from different cycles.
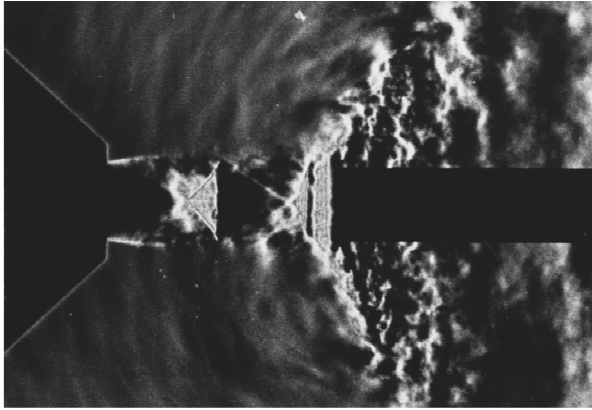
FIG. 8. Schlieren photograph taken at $p_0/p_a = 3.59$ and $h/d = 3.04$ where no tone is produced.

FIG. 9. Schlieren photograph taken when a convergent–divergent nozzle was used at a stagnation pressure of $p_0/p_a = 4.23$ and an impinging plate was located at $h/d = 3.16$. The pressure ratio for ideal expansion is approximately 2.84.

throughout the cycle. The second shock wave oscillates somewhat randomly about a mean position. The downstream location of the standoff shock wave is relatively constant except for its disappearance in two of the photographs. The oscillations of the first shock wave when S1 tones are produced are quite violent and are phase locked with the sound waves [see Figs. 6 and 7 in Henderson and Powell (1996)]. This is in contrast to the slight (random) motion of the shock waves occurring when S2 tones are produced. This is another indication that the generation mechanisms for the S1 and S2 tones are fundamentally different.

The schlieren photograph in Fig. 8 was taken at a stagnation pressure just below where S2 tones occur. By comparing the photographs in Fig. 5(a) and Fig. 8, which were both taken at a nozzle-to-plate spacing of $h/d = 3.04$, it is possible to see that S2 tones are produced when a Mach disk cuts off the apex of the first conical shock wave and the second shock wave is dome shaped rather than conical. The deflection angle of the flow downstream of the plate when no tones are produced is quite large compared to the deflection angle occurring when S2 tones are produced. This may be due to flow separation which may occur when S2 tones are not produced.

Powell (1988) presented timed schlieren photographs for S2 tones, although at that time, no distinction was made between S1 and S2 tones. He found similar results to those presented above; however, it was not possible to see a standoff shock wave in his schlieren photographs and no shadowgraph photographs were included in the study. Because of this, what was considered to be the standoff shock wave oscillating and changing shape was most likely the second shock wave.

IV. EXPERIMENTS USING A CONVERGENT–DIVERGENT NOZZLE

The convergent–divergent nozzle discussed by Henderson and Powell (1996) was used in place of the convergent nozzle in an effort to eliminate all shock waves except the standoff shock wave. The purpose of this experiment was to determine the importance of the first and second shock waves to the production of S2 tones. However, it was discovered that the design pressure for ideal expansion was much higher than the actual pressure for ideal expansion. [See the discussion in Henderson and Powell (1996) for more details]. The schlieren photograph in Fig. 9 shows that a shock wave still exists, but the structure of the jet is quite different from the choked impinging jet at the same operating conditions. At this operating condition, i.e., $p_0/p_a = 4.23$ and $h/d = 3.16$, an S2 tone occurs when the convergent nozzle is used. However, no tone occurs at these operating conditions when the convergent–divergent nozzle is used. This is a further indication that the structure of the jet is important to the production of S2 tones and that altering the jet structure has the potential of eliminating these tones.
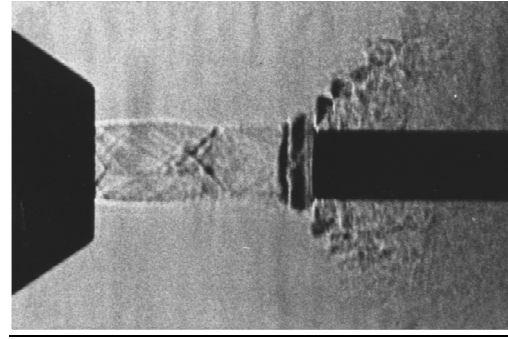
V. REFLECTOR EXPERIMENTS

In the original experiments of Henderson and Powell (1993), reflector tests were performed to determine if the small plate tones were part of a feedback loop to the nozzle. A nozzle extension was placed at the nozzle exit and a large movable reflector with a diameter equal to 35 nozzle diameters was placed behind the nozzle extension exit. The face of the reflector was perpendicular to the jet axis. The results indicated that all small plate tones were part of a feedback loop to the nozzle; however, since it was never possible to eliminate the tones and the reduction in the amplitude of the tones was not significant, the results were to be only preliminary.

The idea behind the use of a reflector is that if it is possible to perfectly cancel the incoming sound waves at the nozzle exit, then tones which are part of a feedback loop to the nozzle will not be produced. The nozzle exit is the most sensitive portion of the jet to the creation of stream disturbances. If the sound creating these disturbances is eliminated, the disturbances must be eliminated, and so the sound resulting from the interaction of these disturbances with obstacles or gradients in the jet will not be produced. Ideally, perfect cancellation would occur when the reflector is infinite in size and is located at $\lambda/4$ or $3\lambda/4$ behind the nozzle exit; however, perfect cancellation is not possible since a nozzle extension extends into the flow and the reflector has a finite size. Since the reflector is quite large compared to the wavelength of sound, it is expected that significant cancellation will still occur at about $\lambda/4$ or $3\lambda/4$.
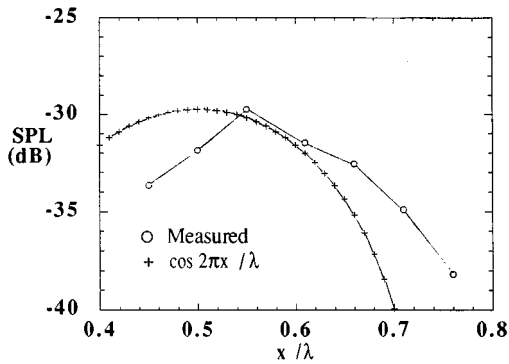
FIG. 10. Far-field sound-pressure levels as a function of the nozzle-to-reflector spacing, x, when $h/d = 3.13$ and $p_0/p_a = 4.20$. The measured and calculated maxima have been made equal.

FIG. 11. Model showing (a) the proposed location of the sound sources in the jet and (b) the distances used in the array model.

The reflector tests of Henderson and Powell (1993) have been performed again with a greater range of nozzle-to-reflector distances, the same nozzle extension, and a new reflector. The details of the new reflector can be found in Henderson and Powell (1996). Figure 10 shows the results of the reflector tests, where x is the nozzle-to-reflector distance. An arbitrary datum has been used for both measured and calculated values. The calculated values in the plot represent the sound-pressure levels at the nozzle exit for the situation of perfect reflection, and the experimental values are the measured sound-pressure levels in the far field at a fixed nozzle-to-plate spacing and angle to the jet. It is expected that, for the case of feedback to the nozzle, the sound-pressure levels in the far field will be related to the sound-pressure levels at the nozzle exit. The maximum measured sound-pressure level in the far field occurs between 0.5λ and 0.6λ, which is close to the expected value of 0.5λ. The tones disappear at 0.8λ, which is close to the expected value of $3\lambda/4$ for perfect reflection. These results strongly indicate that S2 tones are part of a feedback loop to the nozzle, as suggested by Henderson and Powell (1993).

VI. A SOUND-PRODUCTION MODEL

A. Wavelength for maximum reinforcement

The schlieren photographs indicate that the sound waves originate from the region of the second shock wave, the standoff shock wave, and the shock waves in the deflected flow. Since the tones are part of a feedback loop to the nozzle, it is likely that these tones are the result of the interaction of stream disturbances with these shock waves. This is the same type of mechanism which produces choked jet screech [see Powell (1953)]. The interaction of vorticity with a shock wave has been shown to produce sound [see Ribner (1954) and Meadows *et al.* (1989)].

Powell (1962) used a simple source array with the sources located at three or four shock waves in the jet to explain screech tone frequencies and to effectively describe the far-field acoustic radiation produced by choked jets. This same approach was used by Harper-Bourne and Fisher (1974) to describe the far-field acoustic radiation associated with broadband shock associated noise produced by the interaction of turbulence with shock waves in jets operating
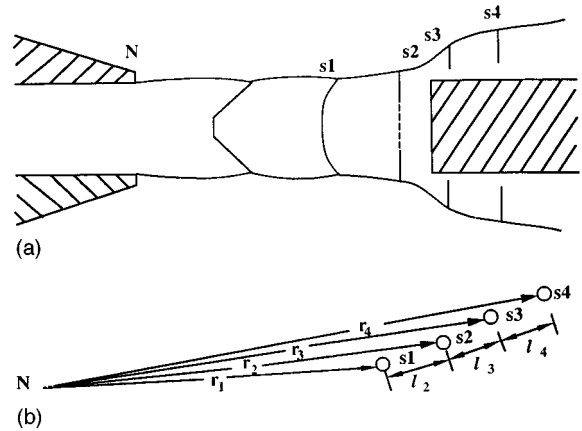
above the critical pressure. Although Rayleigh (1896) showed that the interaction of sound with fluid inhomogenities produces dipole and monople type scattering, and it is expected that convection of vorticity through density changes in the jet may similarly produce monopole and dipole type sound, a simple source array has been chosen to describe the production of S2 tones due to the favorable results obtained by Powell (1962) and Harper-Bourne and Fisher (1974).

Figure 11 shows the location of the sound sources and a representation of the proposed feedback loop. Simple sources have been placed at four locations corresponding to the locations of the second shock wave, the standoff shock wave, and the two stationary shock waves in the deflected flow. Since the tones are part of a feedback loop to the nozzle, it is expected that the wavelength of the discrete tones is the wavelength which produces maximum reinforcement of the pressure waves at the nozzle exit from the four sound sources. This will create the largest amplitude of the disturbances as long as the wavelength is within the region of instability of the jet. For a feedback flow, it is expected that the discrete tones will be related to the amplitude of the stream disturbances.

The pressure due to a monopole is given by

$$p_a(r,t) = \frac{A}{r} e^{i\omega(t - r/c)}, \qquad (1)$$

where r is the distance between the monopole and the observation point, A is the amplitude, ω is the angular frequency, and c is the speed of sound. The distances between the nozzle and source 1 (s1), source 2 (s2), source 3 (s3), and source 4 (s4) are r_1, r_2, r_3, and r_4 respectively (see Fig. 12). Using these distances, the acoustic pressure at the nozzle exit is given by

$$p_N(r,t) = A \sum_{n=1}^{4} \frac{e^{i\omega(t - r_n/c - \tau_n)}}{r_n}, \qquad (2)$$

where τ_n is the time shift at source location n due to the difference in the phase of the disturbance (fundamental frequency only) at source location s1 and source location sn at any instant in time. The value of τ_1 is 0. In Eq. (2), the
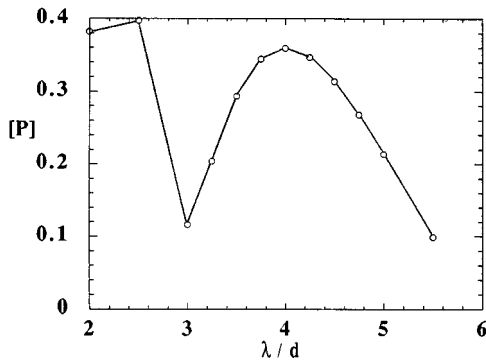
FIG. 12. A plot of the calculated pressure amplitude at the nozzle exit versus the wavelength/d for $h/d=3.04$ and $p_0/p_a=4.06$.

FIG. 13. A schematic of the geometry used for the far-field directivity calculations.

amplitudes of the sources have been assumed to be equal.

The phase shift terms in Eq. (2) may be obtained by assuming a convection speed, u_{con}, of the disturbances in the jet and measuring the distances between the sources in the photographs. The phase shift of each source is then given by l/u_{con}, where l is the distance between the sources as shown in Fig. 11. This allows Eq. (2) to be written as

$$p_N(r,t)=Ae^{i\omega(t-r_1/c)}\sum_{n=1}^{4}\frac{1}{r_n}\exp[-i2\pi(\Delta r_n$$
$$+l_n/M_{ef})/\lambda],\qquad(3)$$

where $\Delta r_n\approx r_n-r_1$, $M_{ef}=u_{con}/c$, $\Delta r_1=0$, $l_1=0$, u_{con} is the convection speed of the disturbances, and λ is the wavelength of sound. The pressure at the nozzle may be represented as $p_N(r,t)=[P]e^{i\omega(t-r1/c+\phi)}$, where $[P]$ is the pressure amplitude at the nozzle exit and is a function of the acoustic wavelength, λ. The value of the $[P]$ may be found from

$$[P]^2=\left[\sum_{n=1}^{4}\frac{1}{r_n}\cos\frac{2\pi}{\lambda}\left(\Delta r_n+\frac{l_n}{M_{ef}}\right)\right]^2$$
$$+\left[\sum_{n=1}^{4}\frac{1}{r_n}\sin\frac{2\pi}{\lambda}\left(\Delta r_n+\frac{l_n}{M_{ef}}\right)\right]^2\qquad(4)$$

once the values of Δr_n, r_n, l_n, and M_{ef} have been determined.

When the value of M_{ef} is taken as 0.4 and the necessary distances are measured from Fig. 6, the plot of $[P]$ vs λ/d shown in Fig. 12 is obtained. In this figure, a peak occurs at a nondimensional wavelength of $\lambda/d=4$ which corresponds to the measured wavelength of the discrete tone produced at this operating condition. This is the expected result for a tone which is part of a feedback loop to the nozzle.

A value of 0.4 for M_{ef} was also used by Powell (1962) in his calculations for the choked jet. The value of the convected Mach number in an ideal shear layer is 0.5; however, shock waves do not occur in the ideal flow and viscous effects are ignored. The experimental determination of the convective Mach number was attempted by measuring the disturbance wavelength in the photographs and comparing this to the measured frequency of sound. This process led to a convective Mach number between 0.2 and 0.3; however,
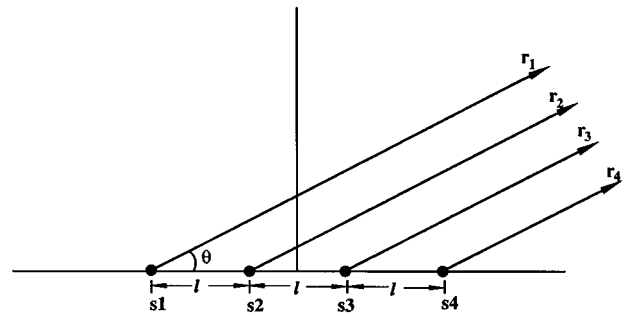
these values are suspect since the disturbance wavelength in the photographs was unclear. Because of this, the value for M_{ef} was taken as 0.4 which appeared to render favorable results in Powell's experiments.

B. Directivity calculations

The schematic in Fig. 13 shows the geometry used for the far-field directivity calculations. Since the directivity measurements are taken in the far field, certain simplifications have been made. The sources have been assumed to be equally spaced and an average value of l has been used. The sources have also been assumed to be located along a straight line. The line connecting the sources is at an angle to the jet axis so the horizontal line joining the sources as shown in Fig. 13 must be tilted (as in Fig. 11) in order to match the measured directivity to the calculated directivity.

The pressure in the far field is given by

$$p_a(r,\theta,t)\approx\frac{A}{r}\left[e^{i(\omega t-kr-3kl\cos\theta/2+3\omega\tau/2)}\right.$$
$$+e^{i(\omega t-kr-kl\cos\theta/2+\omega\tau/2)}$$
$$+e^{i(\omega t-kr+kl\cos\theta/2-\omega\tau/2)}$$
$$\left.+e^{i(\omega t-kr-3kl\cos\theta/2-3\omega\tau/2)}\right],\qquad(5)$$

where $\omega\tau=2\pi l/\lambda M_{ef}$, and k is the wave number. Equation (5) takes the reduced form given by

$$p_a(r,\theta,t)\approx\frac{2A}{r}e^{i(\omega t-kr)}\left\{\cos\left[\frac{3\pi l}{\lambda}\left(\cos\theta-\frac{1}{M_{ef}}\right)\right]\right.$$
$$\left.+\cos\left[\frac{\pi l}{\lambda}\left(\cos\theta-\frac{1}{M_{ef}}\right)\right]\right\}.\qquad(6)$$

The directivity of the sound pressure is, therefore, given by

$$H(\theta)\approx\cos\left[\frac{3\pi l}{\lambda}\left(\cos\theta-\frac{1}{M_{ef}}\right)\right]$$
$$+\cos\left[\frac{\pi l}{\lambda}\left(\cos\theta-\frac{1}{M_{ef}}\right)\right].\qquad(7)$$

The sound-pressure level plot in Fig. 14 may be found using Eq. (7) when the values of M_{ef}, λ/d, and l/d are equal to 0.4, 4, and 0.45, respectively. These values were determined from the photograph in Fig. 6. The maxima of the measured and calculated values have been made equal in
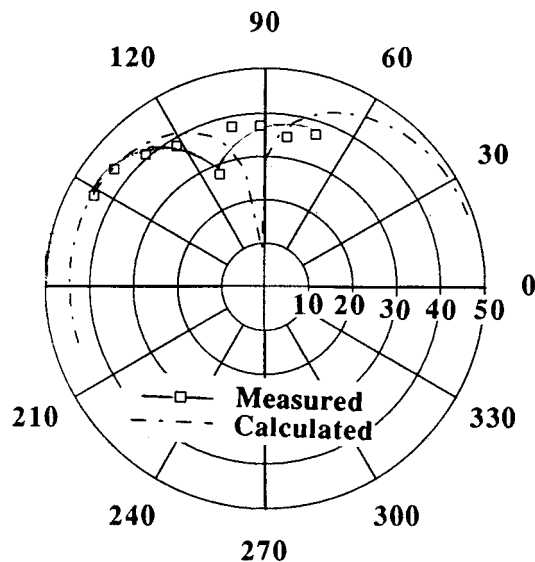
FIG. 14. Calculated and measured directivity using a rotation angle of 20° for the horizontal axis and values of M_{ef}, λ/d, and l/d equal to 0.4, 4, and 0.45 respectively. The experimental data were taken at $h/d=3.04$ and $p_0/p_a=4.06$. An arbitrary datum has been used for both the calculated and measured directivity. The downstream flow direction is at 0°.

this plot. The horizontal axis in Fig. 13 has been rotated 20°, which is the angle between the approximate line connecting the sources and the jet axis in Fig. 6. There are two lobes in the directivity, one upstream and one downstream, with a minimum near 90°. As the angle of the tilted coordinate system increases, the minimum in the calculated values shifts further toward the upstream direction which corresponds to the minimum occurring at a larger angle. The measured directivity is also shown in the plot. The measured directivity also has a maximum in the upstream direction with a decrease in sound-pressure levels near 90°. It is possible to see the beginning of increased downstream radiation but it is not possible to measure the directivity beyond an angle of approximately 70° since the microphone would be in the deflected jet flow. However, the measured pressure begins to increase in the downstream direction as predicted by the array model. There is a slight difference in the angle where the minimum occurs in the measured directivity and the calculated directivity. In addition to this, the minimum for the calculated values is lower than the minimum for the measured values. This will be discussed in the following section. The experimental data, however, display trends which can be explained by the simple source array model.

VII. DISCUSSION

In the above analysis, a comparison is made between the acoustic radiation due to the four element simple source phased array and the measured acoustic radiation. It is possible to explain the upstream radiation, the minimum near 90°, and the beginning of the downstream radiation using the array. There is a slight shift in the minimum near 90° and the measured sound-pressure level at the minimum is not as low as the calculated value. The difference in the measured and calculated values at the minimum is possibly due to the fact

that the flow has slight nonperiodic elements which may cause the exact location of the sources to shift slightly with time. Small shifts in the source locations, the shock waves, will also occur as the vortices travel downstream. This means that the cancellation that occurs near 90° in the calculated directivity may not be realized in the real problem because of the slight shifts in source locations. Another possible source of error in the calculated directivity is the straight line drawn through the sources. The first source does not exactly fall along this line and the slight error associated with the difference between the true location and assumed location of the first source may keep the high degree of cancellation predicted in the model from occurring in the real problem.

The model used to explain the directivity and frequency of the S2 tones consists of point sources located in two dimensions. In reality, the sources are distributed along annuli encircling the jet. In addition to this, the presence of the jet causes refraction effects due to the interaction of the sound waves and the flow. Since the velocity of the jet is different at different axial and radial locations, the refraction effects vary throughout the jet. For supersonic regions in the jet, the sound waves are unable to propagate through the jet and there is total reflection from the jet boundary. However, the reasonable agreement between the calculated and measured frequency and the general trends observed in the directivity studies indicate that the simplified model presented in the previous section is sufficient for the types of calculations considered in the present analysis.

VIII. CONCLUSIONS

When S2 tones are produced, three shock waves are present in the jet upstream of the plate and two annular shock waves occur in the deflected flow downstream of the plate. There does not appear to be a phase lock between the sound waves and the motion of the shock waves in the jet. Photographic evidence indicates that the sound waves originate from the second shock wave, the standoff shock wave, and the shock waves in the deflected flow. The sound is part of a feedback loop to the nozzle and the frequencies of the S2 tones are explained using a four element simple source phased array. The general trends observed in directivity studies for the S2 tones are also explained by the array model.

Didden, N., and Ho, C-M. (**1985**). ''Unsteady separation in a boundary produced by an impinging jet,'' J. Fluid Mech. **160**, 235–256.

Harper-Bourne, M., and Fisher, J. J. (**1974**). ''The noise from shock waves in supersonic jets,'' Noise Mechanisms, AGARD-CP-131, 11.1-11.13.

Hartmann, J., and Trolle, B. (**1927**). ''A new acoustic generator. The air-jet-generator,'' J. Sci. Instrum. **4**, 101–111.

Henderson, B., and Powell, A. (**1993**). ''Experiments concerning tones produced by an axisymmetric choked jet impinging on flat plates,'' J. Sound Vib. **168,** 307–326.

Henderson, B., and Powell, A. (**1996**). ''Sound production mechanisms of the axisymmetric supersonic jet impinging on small plates. The production of primary tones,'' J. Acoust. Soc. Am. **99,** 153–162.

Ho, C-M., and Nosseir, N. S. (**1981**). ''Dynamics of an impinging jet, Part 1. The feedback phenomenon,'' J. Fluid Mech. **105,** 119–142.

Marsh, A. (**1961**). ''Noise measurements around a subsonic air jet impinging on a plane, rigid surface,'' J. Acoust. Soc. Am. **33,** 1065–1066.

Meadows, K., Kumar, A., and Hussaini, M. (**1989**). ''A computational study on the interaction between a vortex and a shock wave,'' AIAA 12th Aeroacoustics Conference, San Antonio, AIAA-89-1043.

Nosseir, N. S., and Ho, C-M. (**1982**). ''Dynamics of an impinging jet. Part 2. The noise generation,'' J. Fluid Mech. **116,** 379–391.

Powell, A. (**1953**). ''On the mechanism of choked jet noise,'' Proc. Phys. Soc. London, Sec. B **66,** 1039–1056.

Powell, A. (**1962**). ''Nature of the feedback mechanism in some fluid flows producing sound,'' Fourth International Congress on Acoustics, Copenhagen.

Powell, A. (**1988**). ''The sound-producing oscillations of round underexpanded jets impinging on normal plates,'' J. Acoust. Soc. Am. **83,** 515–533.

Powell, A., and Henderson, B. (**1990**). ''On the tones of round choked jets impinging on normal flat plates,'' AIAA 13th Aeroacoustics Conference, AIAA-90-3985.

Rayleigh, Lord (**1896**). *The Theory of Sound* (Reprinted by Dover, New York, 1945), 2nd ed., Vol. 2.

Ribner, H. S. (**1954**). ''Convection of a pattern of vorticity through a shock wave,'' National Advisory Committee for Aeronautics, Report 1164.

Sound generation by interaction of two inviscid two-dimensional vortices

S. K. Tang

Department of Building Services Engineering, The Hong Kong Polytechnic University, Hong Kong

N. W. M. Ko[a)]

Department of Mechanical Engineering, The University of Hong Kong, Hong Kong

Sound generated by the pairing of two inviscid finite-core-size two-dimensional vortices is studied numerically using the contour dynamics method and vortex sound theory. Results indicate that the resulting sound field is a lateral quadrupole type and its axis rotates with nonuniform speed about the origin of the source region. Results also show that the accelerating and decelerating motions of the vortices are important in the sound generation. © *1997 Acoustical Society of America.* [S0001-4966(97)03708-9]

PACS numbers: 43.28.Ra, 43.28.Py, 43.50.Nm [LCS]

INTRODUCTION

The theories of Lighthill[1] and Curle[2] show that sound can be generated by unsteady turbulent flow. In the low Mach number case, the results of Powell[3] and Howe[4] show that the unsteady motion of vorticity is important in sound generation. Studies of vortex sound are, therefore, important in understanding the sound generation of flow noise problems. Typical examples of such problems are jet noise,[5] boundary layer noise,[6] and noise generated from helicopters.[7] Scaling laws for turbulent flow sound intensity have been proposed and some comparisons with experimental results have also been done (for instance, Lush[8]). However, studies, which explicitly relate the sound generation mechanism to vortex motion, are very limited, though a detailed discussion on aerodynamic noise generation is given by Crighton.[9]

It is now well known that sound is generated when two vortices interact. In an axisymmetric low Mach number jet, noise is generated by the pairing of vortex rings.[5] Möhring[10] analytically worked out the sound field generated by the pairing of two thin core closely spaced vortex rings, while Shariff *et al.*[11] computed the sound radiated from inviscid thick core vortex ring pairing using the method of contour dynamics. Tang and Ko[12] showed further that the axial jerk and radial acceleration of the vorticity centroids of the vortex rings play an important role in the pairing noise generation.

In a low Mach number two-dimensional mixing layer and rectangular jet of large-aspect ratio, sound is expected to come from the pairing of two-dimensional vortices. As the near-field source is not compact in this case, the far-field noise is often calculated numerically. For simplicity, the two-dimensional vortices are usually modeled as line vortices during analysis. Recently, Lee and Koo[13] used this technique to compute the sound field pattern generated by two spinning rectilinear vortices in the context of incompressible and inviscid flow. Mitchell *et al.*[14] extended this study to compressible flow. However, they have not addressed the

sound generation mechanism involved. In addition, vortices with infinitesimal core sizes are rare in reality. A thick core model appears to be more appropriate for the study and understanding of the mechanism.

In the present study, the sound field generated by the pairing of two equal inviscid thick core two-dimensional vortices is investigated using the method of contour dynamics. Special attention is given to the mechanism of the sound generation. For simplicity, the effect of compressibility is neglected.

I. THEORETICAL AND COMPUTATIONAL CONSIDERATIONS

This section gives brief discussions on the contour dynamics method and vortex sound theory. The two-dimensional pairing noise generation mechanism is also discussed using existing theories.

A. Method of contour dynamics

The shapes of the vortex cores during pairing can be obtained using the method of contour dynamics.[15] This is done by first observing the streamfunction ψ in two dimensions at a point $\mathbf{y}_0$ in the flow field due to a vortex:

$$\psi(\mathbf{y}_0) = -\frac{1}{2\pi} \int \int \omega \ln|\mathbf{y}_0 - \mathbf{y}| dy_1 \, dy_2, \tag{1}$$

where y_1 and y_2 denote the longitudinal and transverse coordinates, respectively (Fig. 1). This integration is done across the cross section of the vortex. As $\mathbf{u} = \nabla \times \hat{y}_3 \psi$, where y_3 represents the spanwise coordinate and the caret denotes unit vector, use of Stoke's theorem gives the contour integral

$$\mathbf{u}(\mathbf{y}_0) = -\frac{1}{4\pi} \int \omega \ln|\mathbf{y}_0 - \mathbf{y}|^2 \, d\mathbf{y} \tag{2}$$

and the singularity in Eq. (2) can be removed by an integration by parts as in Dritschel.[16] Details of the derivation of the method can be found in Zabusky *et al.*[15] and Dritschel.[16] In the present investigation, the node point insertion and relaxation scheme of Pozrikidis and Higdon[17] is used to ensure a

---

[a)]Corresponding author. Electronic mail: NWMKO@hkucc.hku.hk
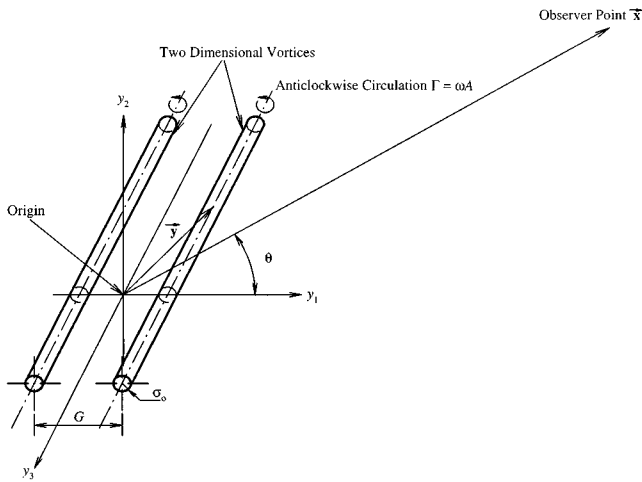
FIG. 1. Schematic diagram of two-dimensional vortices.

smooth contour during the computation. All the contour integrals involved are computed using the four-point Gaussian quadrature procedure and the core shapes are obtained by the fourth-order Runge–Kutta method as in Tang and Ko[12] and Dritschel.[16] The time step is so chosen that no visible difference of results can be found upon its further reduction. The definition of vortex centroid[17] is adopted as

$$\mathbf{y}_c = y_{c1}\hat{y}_1 + y_{c2}\hat{y}_2 = \frac{\int \omega \mathbf{y} \, dA}{\int \omega \, dA},$$ (3)

where $\mathbf{y}_c$ is the position vector of the vortex centroid having the coordinates (y_{c1}, y_{c2}) (see Fig. 2). The velocity, acceleration, and jerk of the vortex in the rest of the paper simply refer to the first, second, and third time derivatives of y_c, respectively. For simplicity, the two vortices are placed symmetrically about the origin at the beginning of the computation. Uniform vorticity within vortex cores is assumed so as to fulfill the requirement of $d\omega/dt = 0$. Under this circumstance, the vortex centroid, as defined by Eq. (3), coincides with the center of mass of the incompressible vortex.

The effect of viscosity on the far-field pressure is not important in the spinning vortex case of Mitchell *et al.*[14] and
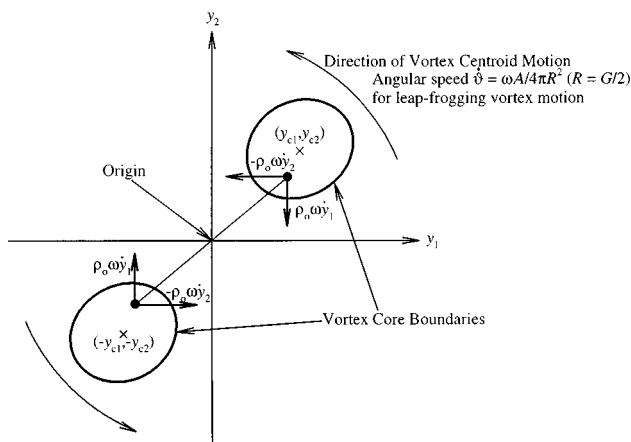
FIG. 2. Forces on core fluids. ×: Vortex centroid.

it is neglected in the present study. Without loss of generality, the vorticity is assumed to be positive (anticlockwise) throughout this study.

B. Vortex sound

The sound field from turbulent flow can be obtained by solving the following inhomogeneous wave equation:

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} - \nabla^2 p = q,$$ (4)

where p, c, and t are the far-field pressure, ambient speed of sound, and observer time, respectively. Here, q is the source term, which in the low Mach number case, can be written as $\rho_0 \nabla \cdot (\omega \times \mathbf{u})$, where ω, ρ_0, and $\mathbf{u}$ are vorticity, ambient fluid density, and the velocity of the vorticity propagation, respectively.[3,4] For convenience, boldfaced symbols denote vector quantities in the rest of the paper. Following the method of Möhring,[10] one obtains for a point source of volume dV

$$dp(\mathbf{x},t) = \frac{\rho_0}{12\pi c^2}\frac{\partial^3}{\partial t^3}\frac{(\hat{\mathbf{x}}\cdot\mathbf{y})\mathbf{y}\cdot(\omega\times\hat{\mathbf{x}})}{|\mathbf{x}-\mathbf{y}|}\,dV,$$ (5)

where the expression being differentiated is to be evaluated at retarded time $t - |\mathbf{x}-\mathbf{y}|/c$. Here, $\mathbf{x}$ and $\mathbf{y}$ denote, respectively, far- and source-field distances (Fig. 1) and $\hat{\mathbf{x}}$ is an unit vector in the direction $\mathbf{x}$ and is nondimensional. For compact vortex cores, the contribution to the far-field noise from an infinitesimal length dy_3 at $\mathbf{y}$ along a two-dimensional vortex of cross section area A is

$$dp(\mathbf{x},t) = \frac{\rho_0\,dy_3}{12\pi c^2}\frac{\partial^3}{\partial t^3}\left[\frac{1}{|\mathbf{x}-\mathbf{y}|}\left(\cos(2\theta)\int \omega y_1 y_2\,dA\right.\right.$$
$$\left.\left.+ \sin(2\theta)\int \frac{y_2^2 - y_1^2}{2}\,\omega\,dA\right)\right],$$ (6)

where θ is the observer angle (Fig. 1). The total far-field pressure $p(\mathbf{x},t)$ generated by the vortices can be obtained by integrating Eq. (6) along y_3:

$$p(\mathbf{x},t) = \frac{\rho_0}{12\pi c^2}\int \frac{1}{|\mathbf{x}-\mathbf{y}|}\frac{\partial^3}{\partial t^3}\left(\cos(2\theta)\int \omega y_1 y_2\,dA\right.$$
$$\left.+ \sin(2\theta)\int \frac{y_2^2 - y_1^2}{2}\,\omega\,dA\right)dy_3.$$ (7)

Equation (7) suggests the sound field is a lateral quadrupole. The constant outside the integral in Eq. (7) has no bearing on the sound generation process and is thus ignored in the foregoing discussions. For easy reference, S_1, S_2, D_1, and D_2 are used in the rest of the paper to denote the following integrals:

$$S_1 = \int \omega y_1 y_2\,dA, \quad S_2 = \frac{1}{2}\int \omega(y_2^2 - y_1^2)dA,$$

$$D_1 = \int \frac{\dddot{S_1}}{|\mathbf{x}-\mathbf{y}|}\,dy_3, \quad D_2 = \int \frac{\dddot{S_2}}{|\mathbf{x}-\mathbf{y}|}\,dy_3.$$ (8)

S_1 and S_2 can be computed using contour integrals and Stoke's theorem. The triple overdot denotes triple time de-

rivative. D_1 and D_2, which are referred as the far-field quadrupole strengths in the rest of the paper, are calculated using the procedure of Ffowcs Williams and Hawkings.[18] They observed that these integrals can be transformed into time integrals by replacing y_3 by a variable t' and because of the singularity at $t' = t - r/c$, one obtains

$$D_i = \int_{-\infty}^{\infty} \frac{\dddot{S}_i(\mathbf{y}, t - |\mathbf{x} - \mathbf{y}|/c)}{|\mathbf{x} - \mathbf{y}|} dy_3$$

$$= \int_{-\infty}^{t-r/c} \frac{2c\dddot{S}_i(\mathbf{y}, t')}{\sqrt{c^2(t-t')^2 - r^2}} dt'$$

$$\approx \sqrt{\frac{2}{r}} \int_{-\infty}^{t-r/c} \frac{c\dddot{S}_i(\mathbf{y}, t')}{\sqrt{c(t-t') - r}} dt', \qquad (9)$$

where $r^2 = (x_1^2 - y_{c1}^2) + (x_2^2 - y_{c2}^2)$ and $i = 1,2$. It should be noted that S_i, are independent of y_3. Equation (9) can be solved numerically by further substituting $c(t - t') - r$ by β^2 with the assumption that there is no sound at $\mathbf{x}$ for $t < r/c$. Therefore, the sound field can be solved provided that the time variations of the terms $\dddot{S}_i$ can be obtained. D_i are also the strengths of the two-dimensional far-field quadrupoles as shown in Eq. (7), while $\dddot{S}_i$ denotes the source strengths [Eq. (8)].

C. Sound generation mechanism

Though Eq. (9) suggests that the quadrupole strength depends on D_1 and D_2, the sound generation mechanism involved is more related to S_1 and S_2. The appearance of the former two integrals are only due to the noncompactness of the source region and thus have nothing to do with the vortex motions, not to say the sound generation mechanism. The sound generation mechanism is therefore discussed with Eq. (6).

Under the assumptions of constant vorticity and incompressible vortex, Eq. (6) shows that the sound radiated to the far field at an observer angle θ equals the third time derivative of the cross product of inertia in the angle $(180° - \theta)$,[19] showing that the deformation of the vortex core is important in the sound generation process. However, it is more realistic to describe the sound generation mechanism in terms of some easily quantified parameters, such as force and acceleration of the vortex centroid,[12] as they are easier to verify experimentally than the core shapes. The following paragraphs show how $\dddot{S}_i$ can be related to the motion of the vortex centroids and describe the physical meanings of the terms involved.

Allowing for nonuniform vorticity distribution within the vortex core, the impulse $\mathbf{I}$ of an individual vortex with finite core size is given by the expression

$$\mathbf{I} = \rho_0 \int \mathbf{y} \times \omega \, dA = -\rho_0 \hat{y}_1 \int y_2 \omega \, dA$$

$$+ \rho_0 \hat{y}_2 \int y_1 \omega \, dA. \qquad (10)$$

As the vortices are mirror images of each other, the total impulse of the interacting vortices vanishes but those of the

individual vortices are vector quantities in two-dimensional space. The force $\mathbf{F}$ acting on each vortex equals the rate of change of impulse of the vortex and thus

$$\mathbf{F} = \frac{d\mathbf{I}}{dt} = -\rho_0 \hat{y}_1 \int \dot{y}_2 \omega \, dA + \rho_0 \hat{y}_2 \int \dot{y}_1 \omega \, dA, \qquad (11)$$

where the overdot represents differentiation with respect to time. The forces acting on a small mass of fluid within the vortex cores are illustrated in Fig. 2. Equations (10) and (11) show that the total force acting on the core fluids can be estimated by finding the velocity of the vortex centroids when ω is constant. Also, the conservation of angular momentum J requires that

$$\frac{dJ}{dt} = 0 \Rightarrow \frac{d}{dt} \int \omega(y_1^2 + y_2^2) dA = 0. \qquad (12)$$

Besides, the sound energy radiated is negligible compared to the flow energy, especially in the low Mach number case.[1] The total kinetic energy per unit length of the two-dimensional flow K remains fairly constant and thus, following the work of Lamb,[20]

$$\frac{dK}{dt} = 0 \Rightarrow \rho_0 \frac{d}{dt} \int (\dot{y}_1 y_2 - \dot{y}_2 y_1) \omega \, dA = 0. \qquad (13)$$

Recalling Eq. (6), one obtains together with Eqs. (12) and (13) that

$$\frac{\partial^3 S_1}{\partial t^3} = \frac{\partial^3}{\partial t^3} \int \omega y_1 y_2 \, dA$$

$$= \frac{\partial^2}{\partial t^2} \int \omega(y_1 \dot{y}_2 + y_2 \dot{y}_1) dA$$

$$= \frac{\partial^2}{\partial t^2} \int \omega(2y_1 \dot{y}_2 + y_2 \dot{y}_1 - y_1 \dot{y}_2) dA$$

$$= \frac{\partial^2}{\partial t^2} \int 2\omega y_1 \dot{y}_2 \, dA \qquad (14a)$$

and

$$\frac{\partial^3 S_2}{\partial t^3} = \frac{\partial^3}{\partial t^3} \int \frac{\omega(y_2^2 - y_1^2)}{2} dA$$

$$= \frac{\partial^3}{\partial t^3} \int \frac{\omega(y_2^2 + y_1^2 - 2y_1^2)}{2} dA$$

$$= -\frac{\partial^2}{\partial t^2} \int 2\omega y_1 \dot{y}_1 \, dA. \qquad (14b)$$

Similarly, one can also show that

$$\frac{\partial^3 S_1}{\partial t^3} = \frac{\partial^2}{\partial t^2} \int 2\omega y_2 \dot{y}_1 \, dA$$

and $\qquad (14c)$

$$\frac{\partial^3 S_2}{\partial t^3} = \frac{\partial^2}{\partial t^2} \int 2\omega y_2 \dot{y}_2 \, dA.$$

It is observed from Eqs. (10) and (14) that $\partial^3 S_2 / \partial t^3$ represents the total anticlockwise couple about the origin acting on the vortex core fluids (Fig. 2). It is therefore related to the

rotation of the vortex core fluids about the origin. However, one should bear in mind that the total couple acting on the core fluids vanishes because the anticlockwise and clockwise couple balance each other in the present case.

The meaning of $\partial^3 S_1 / \partial t^3$ is far less trivial. It is the magnitude of the product of a force and a distance in one direction; either in the y_1 or y_2 direction. As the vortex centroids are rotating anticlockwise about the origin ($\omega > 0$), Eq. (11) shows that the total force acting on one vortex core is always pointing towards the two axes (Fig. 2). Here, $\partial^3 S_1 / \partial t^3$ seems to represent somewhat like the "work done" in one direction against the displacement of the vortex core fluids from the axis in that direction. Thus sound is generated by the force acting on individual vortex. However, in this two dimensional case, Eq. (14) cannot be further decomposed to relate the sound generation mechanism with the acceleration and jerk of the vortex centroids as explicitly as in the axisymmetric vortex ring case of Tang and Ko,[12] unless the cores are very thin.

In the thin core case, the vortices undergo circular motion with angular speed $\dot{\vartheta} = \omega A / 4\pi R^2$, where R is the radius of the circular orbit, which equals to half the separation between vortex centroids G.[3] Thus one can observe that

$$y_{c1} = R \cos(\dot{\vartheta} t) \quad \text{and} \quad y_{c2} = R \sin(\dot{\vartheta} t). \tag{15}$$

For core of infinitesimal size, its deformation is negligible and it can be assumed that the vortex cores remain circular throughout the interaction, Eq. (14) can then be simplified to

$$\begin{aligned}
\frac{\partial^3 S_1}{\partial t^3} &= -16 A R^2 \rho_0 (R \dot{\vartheta}^2)^2 \cos(2\dot{\vartheta} t) \\
&= 16 A R^2 \rho_0 [(\ddot{y}_{c2})^2 - (\ddot{y}_{c1})^2] = E_1, \\
\frac{\partial^3 S_2}{\partial t^3} &= -16 A R^2 \rho_0 (M \dot{\vartheta}^2)^2 \sin(2\dot{\vartheta} t) \\
&= -32 A R^2 \rho_0 \ddot{y}_{c1} \ddot{y}_{c2} = E_2,
\end{aligned} \tag{16}$$

where E_i represent the source strengths associated with non-linear coupling of longitudinal and transverse accelerations. Equations (16) also show the importance of the accelerations $\ddot{y}_{c1}$ and $\ddot{y}_{c2}$ in the sound generation in the thin core case. The constant ambient density ρ_0 will be neglected in the rest of the discussion. It should be noted in this nondeformable thin vortex core case that the accelerations $\ddot{y}_{c1}$ and $\ddot{y}_{c2}$ are proportional to the angular speed $\dot{\vartheta}$ and the angular acceleration $\ddot{\vartheta}$ vanishes. When the vortex cores are deformable, $\ddot{y}_{c1}$ and $\ddot{y}_{c2}$ will no longer be proportional to $\dot{\vartheta}$ and $\ddot{\vartheta}$ is time dependent.

The situation in the thick core case is complicated because the vortex core shapes are noncircular and change with time. The angular velocity $\dot{\vartheta}$ and R, therefore, are functions of time so that Eqs. (16) alone are unable to give the sound field. However, R and $\dot{\vartheta}$ for the two vortices are the same as far as their circulations are equal. When the vortex cores are thick, $\dddot{S}_i$ cannot be expressed analytically as in Eqs. (16). It will be shown numerically in the next section that

$$\begin{aligned}
\dddot{S}_1 &\approx E_1 + k_1 (\dot{y}_{c1} \dot{y}_{c2} r_c \ddot{\theta}_c + \tfrac{1}{2}(\dot{y}_{c1}^2 - \dot{y}_{c2}^2) \ddot{r}_c), \\
\dddot{S}_2 &\approx E_2 + k_2 (\dot{y}_{c1} \dot{y}_{c2} \ddot{r}_c - \tfrac{1}{2}(\dot{y}_{c1}^2 - \dot{y}_{c2}^2) r_c \ddot{\theta}_c),
\end{aligned} \tag{17}$$

when the vortices do not coalesce. Here, k_i are constants depending on the separation of the vortex centroids and r_c and θ_c are, respectively, the radial and angular positions of vortex centroid. Therefore, though Eqs. (16) are obtained due to the circular motion characteristics of thin core vortices, E_1 and E_2 are also important in thick core vortex pairing noise generation as far as the vortices do not coalesce. As will further be shown later, numerical data in the case of coalescence favor the following relationships:

$$\begin{aligned}
\dddot{S}_1 &\approx \exp(-\omega(t - r/c)/b)(l_{11} \ddot{y}_{c1} + l_{12} \ddot{y}_{c2}), \\
\dddot{S}_2 &\approx \exp(-\omega(t - r/c)/b)(l_{21} \ddot{y}_{c1} + l_{22} \ddot{y}_{c2}),
\end{aligned} \tag{18}$$

where b and l_{ij} are constants. Thus the sound generation mechanism in the pairing of two-dimensional thick core vortices involves the accelerating or decelerating motions of the vortex core fluids. The numerical results and the derivation of Eqs. (17) and (18) are given in the next section.

II. NUMERICAL RESULTS

The computations were done on the IBM 9076 Scalable POWERparallel system of The University of Hong Kong. Two types of interactions are considered in this study. The first one is the leapfrogging motion and the other the vortex coalescence.[15] In the former type of interaction, the vortex centroids undergo circular motion. Deformation of the vortex cores occurs at the same time but the degree of deformation is not severe. The latter type of interaction involves severe core deformation together with the formation of strips of vorticity around the interacting vortices. These evolutions of the vortex cores are well known[15,21] and it will not be discussed in detail in the present paper. In the following sections, both the sound field and sound generation mechanism are addressed. In order to differentiate observer time and interaction time, t_i represents vortex interaction time which can also be considered as the sound generation time. Numerically, t_i equals the retarded time, that is $t_i = t - r/c$. It is introduced to differentiate the time scale of vortex interaction and of far-field noise radiation. Also, all quantities, unless otherwise stated, are nondimensionalized in the rest of the paper. Lengths, such as y_{c1} and R, are normalized by σ_0, where σ_0 is the initial core radius (Fig. 1), and times, such as t_i and t, by ω^{-1}. Other quantities presented are normalized by combinations of σ_0 and ω. All velocities and accelerations, except the angular velocity and acceleration, which are normalized by ω and ω^2, respectively, are normalized by $\sigma_0 \omega$ and $\sigma_0 \omega^2$, respectively. The terms $\dddot{S}_1$, $\dddot{S}_2$, E_1, and E_2 are normalized by $\sigma_0^4 \omega^4$, and the quadrupole strengths D_1 and D_2 by $\sigma_0^3 \omega^4$.

A. Leapfrogging case

Figure 3 shows the shapes of the initially trailing vortex core for an initial spacing of $G/\sigma_0 = 5$ and angular momentum $J = 21.2$ at time $\omega t_i = 20, 40,$ and 60. Leapfrogging motion without severe core formation is observed. The maxi-
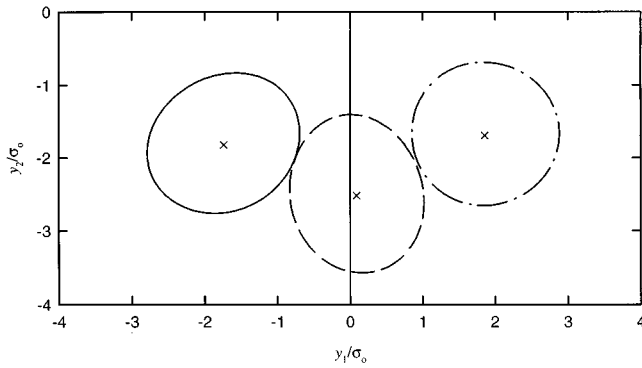
FIG. 3. Variation of initially trailing vortex core shape during interaction. $G/\sigma_0 = 5$, $J = 21.2$. ———: $\omega t_i = 20$; ——-: $\omega t_i = 40$; —·—: $\omega t_i = 60$. ×: Vortex centroid.

mum variations of the cross-sectional area and angular momentum of individual vortices are less than 0.1% and thus it can be assumed that the angular momentum and circulation are conserved. The corresponding time variations of D_1 and D_2 are shown in Fig. 4(a). Wavy oscillations appear on these time variations but the differences between D_1 and D_2 are not only in magnitude but also in phase. As $\dddot{S}_1$ and $\dddot{S}_2$, and
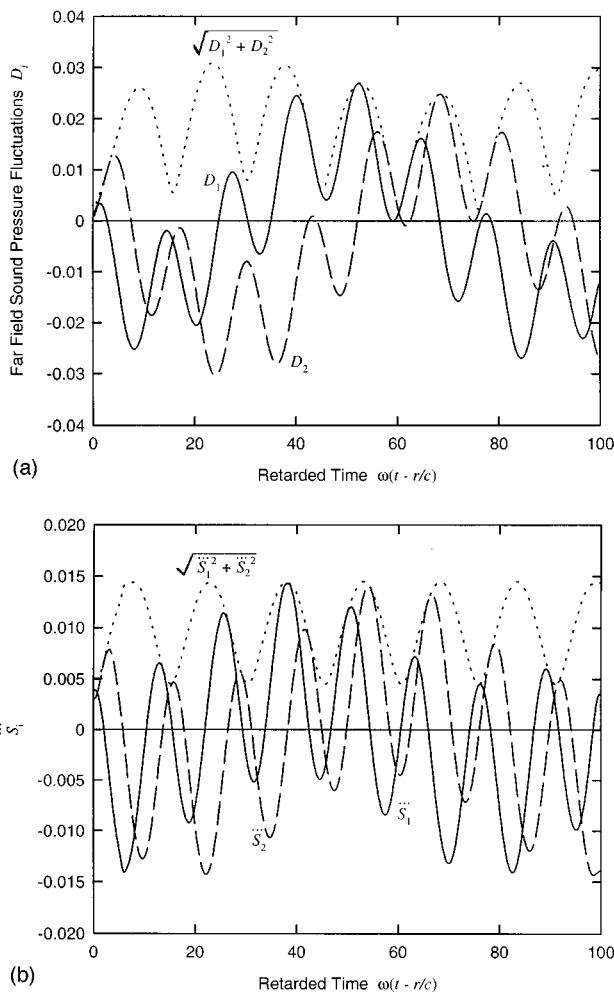
FIG. 4. (a) Time variations of far-field pressure fluctuations during leapfrogging vortex motion. ———: D_1; ——: D_2; ·····: $\sqrt{D_1^2 + D_2^2}$. (b) Time variations of $\dddot{S}_i$. ———: $\dddot{S}_1$; ——: $\dddot{S}_2$; ·····: $\sqrt{\dddot{S}_1^2 + \dddot{S}_2^2}$.

thus D_1 and D_2, represent the magnitudes of two independent quadrupoles as shown in Eq. (7), direct summation of these quantities has no physical meaning. The magnitude of the resultant quadrupole $\sqrt{D_1^2 + D_2^2}$ also fluctuates with time [Fig. 4(a)]. This shows that the vortex system does not radiate constant sound power. Such radiation depends on the states of the vortices as discussed in the next paragraph. It can be noted that the means of D_1 shown in Fig. 4(a) do not vanish, while that for a far field of any source should be zero. The means are about 9% of the corresponding peak values. It is probably due to the time frame shown in Fig. 4(a) which does not cover a full period of the low-frequency component in the computed far-field sound.

As $\dddot{S}_1$ and $\dddot{S}_2$ are more related to the generation of sound, their time variations give information on the sound generation process. The corresponding time variations of $\dddot{S}_1$ and $\dddot{S}_2$ are shown in Fig. 4(b). Similar to Fig. 4(a), the means of $\dddot{S}_1$ appear to be nonzero, probably for the same reason as stated before. The mean values in Fig. 4(b) amount to 5% of the corresponding peak values. It can be observed that these time fluctuations contain low- and high-frequency components. The high-frequency wavy oscillations on these time traces have higher magnitudes relative to those of the corresponding low-frequency components than those on D_1 and D_2. This is probably due to the phase difference of the sound waves radiated to the far-field position $\mathbf{x}$ from different elements along the lengths of the vortices, which tends to smooth out large fluctuations. Both the magnitudes of $\dddot{S}_1$ and $\dddot{S}_2$ vary with ω^3, while those of D_1 and D_2 with $\omega^{3.5}$ (not presented here). The latter is consistent with the existing two-dimensional vortex sound theory.[18] $\dddot{S}_1$ and $\dddot{S}_2$ will be further discussed as they are more related to the mechanism of sound generation. The constant phase shift between $\dddot{S}_1$ and $\dddot{S}_2$, and thus between D_1 and D_2, suggests a lateral quadrupole sound field. In addition, the dotted line in Fig. 4(b) illustrates that the amplitude of sound generated by the interaction attains local minima at $\omega(t-r/c) = \omega t_i = 15.2, 30.4, 45.6, \ldots$, while local maxima are observed at $\omega t_i = 7.7, 22.8, 37.7, \ldots$. At these instants, the vortex system is in the most or least efficient state of sound generation. Defining the differential radius $\Delta\sigma$ as the difference between maximum and minimum distances of the elements on the core boundary from the vortex centroid:

$$\Delta\sigma = [\max(|\mathbf{y} - \mathbf{y}_c|) - \min(|\mathbf{y} - \mathbf{y}_c|)]/\sigma_0,$$

where $\mathbf{y}$ is on the boundary of the vortex core, so that $\Delta\sigma = 0$ corresponds to a circular core, Fig. 5 shows that the minimum $\sqrt{\dddot{S}_1^2 + \dddot{S}_2^2}$ appears at the instant of small differential radius, suggesting that the vortex system radiates less sound energy when the core shape is close to circular. The degree of deviation of the core from circular shape governs the amplitude of sound generation. However, such deviation cannot be easily quantified.

Figure 6 gives the polar plots of normalized far-field sound-pressure amplitudes $|p|$ obtained from Eq. (7) at $\omega t_i = 20$, 40, and 60. The sound fields are lateral quadrupoles. The axes of the quadrupole rotate with nonuniform angular speed about the origin. The major direction of rotation of the
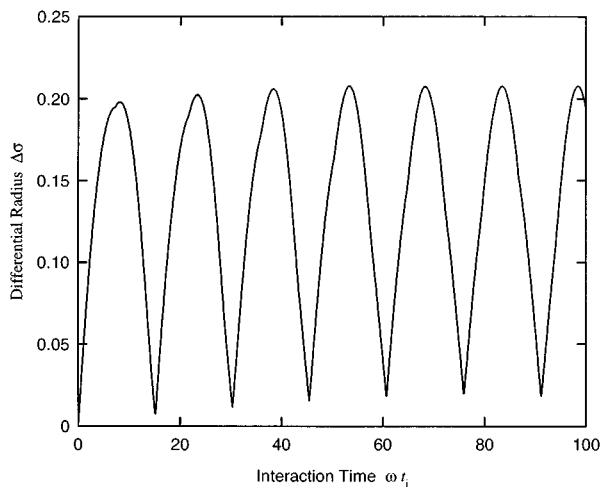
FIG. 5. Time variation of differential radius. $G/\sigma_0 = 5, J = 21.2$.

FIG. 7. Time variation of angle of minimum far-field sound pressure. $G/\sigma_0=5$, $J=21.2$.

quadrupole is anticlockwise and is the same as that of the vortex field. However, there are instants that the quadrupole is rotating clockwise (Fig. 7). This is different from the thin core vortex pairing results of Powell,[3] Lee and Koo,[13] and Mitchell *et al.*[14] which show that the quadrupole always rotates in the direction of the vortex field. The frequency of the wavy oscillation in $\dddot{S}_1$ and $\dddot{S}_2$ [Fig. 4(b)] is not the same as that in Fig. 7. This will be discussed later.

Though the vortex centroids undergo the type of circular motion described by Eqs. (15) with a mean anticlockwise angular velocity $\dot{\vartheta} = \omega A/4\pi R^2$ and their longitudinal and transverse velocities vary nearly sinusoidally with time (Fig. 8), the corresponding accelerations are not (Fig. 9). This may be due to the nutation or rotation of the deformed vortex core as in the vortex ring pairing case.[11,12] However, the low-frequency time variations of these time traces coincide with those calculated using Eqs. (16) which is developed using a nondeformable thin core assumption. The high-frequency wavy oscillations on the acceleration time traces, which have small magnitudes, seem to come from the fluctuating radial and angular accelerations of the vortex centroids in this thick

core case as indicated in Fig. 9. These variations of the radial and tangential accelerations of the vortex centroid, denoted by $\ddot{r}_c$ and $r_c\ddot{\theta}_c$, respectively, are roughly sinusoidal (Fig. 9). For thick core vortices, the high-frequency wavy oscillations on $\dddot{S}_1$ and $\dddot{S}_2$ [Fig. 4(b)] are not expected to come from the linear coupling of velocity and acceleration. Though the magnitudes of the high-frequency wavy oscillations on the radial and tangential acceleration time traces are small (Fig. 9), it will be shown in the next two paragraphs that these oscillations are related to the sound generation process.

The following discussion on the sound generation mechanism for thick core vortex pairing starts from using Eqs. (16). Figures 4(b) and 10 show that the low-frequency components in the source strengths of $\dddot{S}_1$ and $\dddot{S}_2$ can be approximated by E_1 and E_2 given by Eqs. (16), respectively. The time variations of the differences $(\dddot{S}_1 - E_1)$ and $(\dddot{S}_2 - E_2)$ are sinusoidal with period $\omega T = 12.5$ (Fig. 10),

FIG. 6. Polar plot of far-field sound-pressure amplitude. $G/\sigma_0=5$, $J=21.2$. ———: $\omega t_i=20$; ---: $\omega t_i=40$; —·—: $\omega t_i=60$.
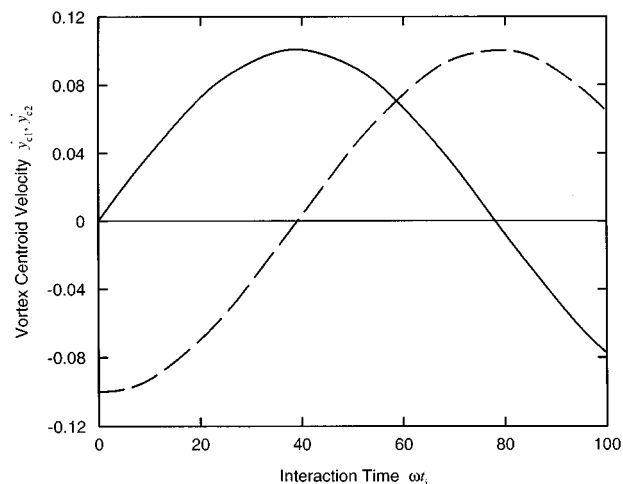
FIG. 8. Time variations of velocities of initially trailing vortex centroid. $G/\sigma_0=5$, $J=21.2$. ———: Longitudinal velocity $\dot{y}_{c1}$; --: transverse $\dot{y}_{c2}$.
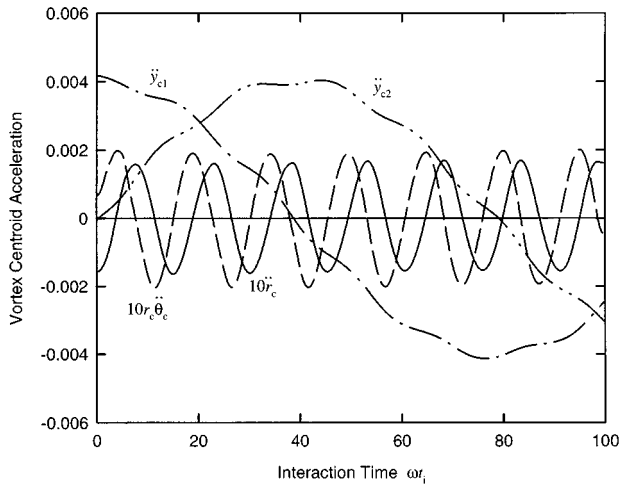
FIG. 9. Time variations of initially trailing vortex centroid accelerations. $G/\sigma_0 = 5$, $J = 21.2$. ———: Radial acceleration $\ddot{r}_c$; – –: tangential acceleration $r_c\ddot{\theta}_c$; —·—: longitudinal acceleration $\ddot{y}_{c1}$; —··—: transverse acceleration $\ddot{y}_{c2}$.

which agrees with that of $\dddot{S}_i$ [Fig. 4(b)]. However, it is different from that of the radial or tangential accelerations shown in Fig. 9, indicating that these wavy oscillations are not generated by these two accelerations alone. Some couplings with other dynamical parameters are expected.

As discussed before, it is more rewarding to relate the sound generation mechanism to the physical parameters that describe the motion of the vortices. From the periods of the high-frequency wavy oscillations of $\dddot{S}_1$ and $\dddot{S}_2$ [$\omega T_s \approx 12.5$, Figs. 4(b) and 10], $\ddot{r}_c$ and $r_c\ddot{\theta}_c$ ($\omega T_a \approx 15$, Fig. 9), and the time variations of $\dot{y}_{c1}$ and $\dot{y}_{c2}$ ($\omega T_v \approx 157$, Fig. 8), which contain negligible high-frequency components, it is found that

$$\frac{1}{T_s} \approx \frac{1}{T_a} + \frac{2}{T_v}.$$

This suggests that the wavy oscillations on $\dddot{S}_1$ and $\dddot{S}_2$ are related to the coupling between the radial/tangential acceleration and the second order of longitudinal or radial veloc-
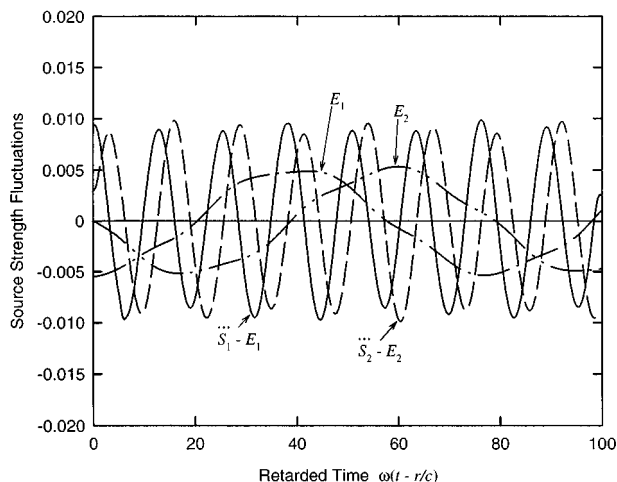
FIG. 10. Time variations of source strengths. $G/\sigma_0 = 5$, $J = 21.2$. ———: $\dddot{S}_1 - E_1$; – –: $\dddot{S}_2 - E_2$; —·—: E_1; —··—: E_2.
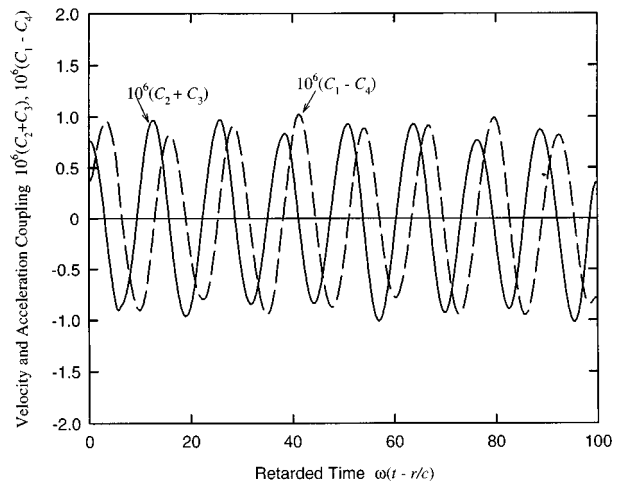
FIG. 11. Time variations of couplings of velocities and accelerations. $G/\sigma_0 = 5$, $J = 21.2$. ———: $(C_2 + C_3)$; – –: $(C_1 - C_4)$.

ity. The Coriolis acceleration component $2\dot{r}_c\dot{\theta}_c$, which is not shown here, is not relevant as its frequency is double those of the acceleration terms. From the sinusoidal characteristics of these acceleration and velocity terms, four types of couplings can be defined. They are named as C_i, where $i = 1, 2, 3$, and 4, as follows:

$$C_1 = \dot{y}_{c1}\dot{y}_{c2}\ddot{r}_c, \quad C_2 = \dot{y}_{c1}\dot{y}_{c2}r_c\ddot{\theta}_c,$$
$$C_3 = \tfrac{1}{2}(\dot{y}_{c1}^2 - \dot{y}_{c2}^2)\ddot{r}_c, \quad C_4 = \tfrac{1}{2}(\dot{y}_{c1}^2 - \dot{y}_{c2}^2)r_c\ddot{\theta}_c. \quad (19)$$

Figures 10 and 11 show that the time variations of $(C_3 + C_2)$ and $(C_1 - C_4)$ follow closely those of $(\dddot{S}_1 - E_1)$ and $(\dddot{S}_2 - E_2)$, respectively. Therefore, the following relationships are suggested:

$$\dddot{S}_1 \approx E_1 + k_1(C_3 + C_2) \quad \text{and} \quad \dddot{S}_2 \approx E_2 + k_2(C_1 - C_4). \quad (20)$$

Here, k_1 and k_2 are dimensional constants with unit m^3 but they are effectively the same (not shown here) and are denoted by k in the rest of the discussion. Equations (20) suggest that the pairing sound is related to the nonlinear couplings between accelerations (both longitudinal, transverse, radial, and tangential) and longitudinal and transverse velocities of vortex centroids. Similar phenomenon is observed for other values of initial spacing G/σ_0 as long as vortex coalescence does not occur. However, Eqs. (20) give better approximation as G/σ_0 increases. The value of k can be estimated using the method of least squares. Figure 12 illustrates the error of sound intensity estimation ϵ that is inherent in the approximation for the first expression in Eqs. (20). This error is defined by the expression

$$\epsilon = \int (\dddot{S}_1 - E_1 - k(C_3 + C_2))^2 \, dt \Big/ \int (\dddot{S}_1)^2 \, dt \quad (21)$$

and is thus a measure of the accuracy of using Eqs. (20) for the approximation of source strengths. It is generally less than 12%. The error associated with $\dddot{S}_2$ is very close to that calculated by the above expression and is not discussed. The error increases with decreasing G/σ_0, but the intensity associated with the high frequency wavy oscillations becomes
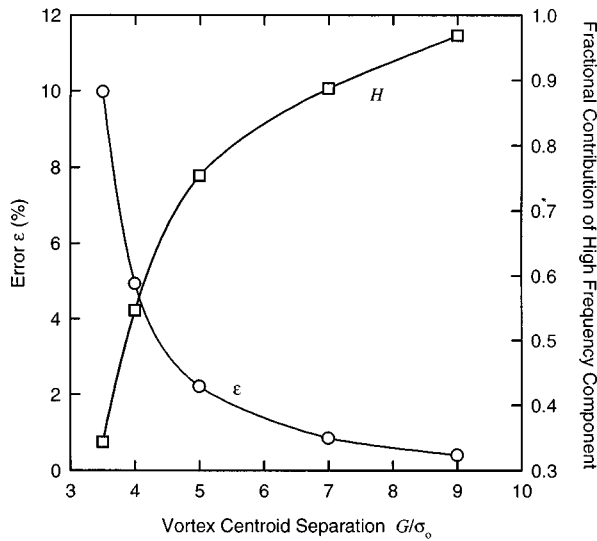
FIG. 12. Error in sound intensity estimation and contribution of high-frequency components in sound radiation. $G/\sigma_0 = 5$, $J = 21.2$. $\bigcirc$: Error in sound intensity estimation by Eq. (21); $\square$: contribution of high-frequency components in overall sound intensity radiation.

FIG. 13. Variation of proportionity constant k with vortex centroid separation for leapfrogging vortex motion. $\bigcirc$: Numerical data from least-squares method; ———: Eq. (23).

less important for small G/σ_0 (Fig. 12). Results for $G/\sigma_0 = 3$ will be discussed in the next section, as coalescence occurs at this separation. Defining the fractional contribution of the high-frequency component in the source strength, which reveals the importance of the high-frequency component in the sound generation, as

$$ H = \frac{\int k^2 (C_3 + C_2)^2 \, dt}{\int (E_1 + k(C_3 + C_2))^2 \, dt}, \tag{22} $$

it can be observed in Fig. 12 that H increases with G/σ_0 and $H \to 1$ as G/σ_0 becomes large. This shows that the couplings between the radial and tangential accelerations and the longitudinal and transverse velocities are important sources of sound for large G/σ_0 but they become insignificant at small G/σ_0, implying that they are not the main source of sound in the coalescence case. This will be discussed later.

Figure 13 shows that $k \propto (G/\sigma_0)^5$. It can also be estimated from Fig. 13 that

$$ k \approx \frac{10}{\pi} \left(\frac{G}{\sigma_0} \right)^5. \tag{23} $$

Equation (23) is included in Fig. 13 and it remains unchanged for other values of ω unless vortex coalescence occurs.

This section shows that the sound generated by the leapfrogging motion of two identical inviscid finite-core-size two-dimensional vortices consists of low-frequency and high-frequency components (Fig. 10). The former is due to the longitudinal and transverse accelerating motions of the vortices, while the latter comes from the coupling of the longitudinal and transverse velocities and the radial and tangential accelerations resulting from the vortex core deformation, showing the importance of the accelerations of vortex centroids in the sound generation mechanism (Fig. 11). Jerk does not appear to be important in this case and this is different from the case of vortex ring pairing.[12]

B. Vortex coalescence case

Figure 14 illustrates the core shapes of the coalescing vortices for $G/\sigma_0 = 3$, $J = 8.36$ at $\omega t_i = 30$. The shapes of the two vortex cores are mirror images of each other. The changes in vortex cross-sectional area and angular momentum are again negligible. Serious core deformation is observed and the motion of the vortex centroids is far from circular, as shown in Fig. 15. Also shown in Fig. 15 is that the vortex centroid does not always rotate in anticlockwise direction, resulting in large changes in its angular velocity at the turning point (Fig. 16). Large changes in its radial and angular accelerations are also observed at this instant (Fig. 17). There is a general tendency for decreasing vortex centroid spatial separation at increased interaction time as shown in Fig. 15, which is also observed in vortex ring coalescence.[12] This decrease in vortex centroid separation is accompanied by an increase in the magnitudes of the longitudinal and transverse accelerations (Fig. 17).

The time variations of the source strengths $\dddot{S}_1$ and $\dddot{S}_2$ are
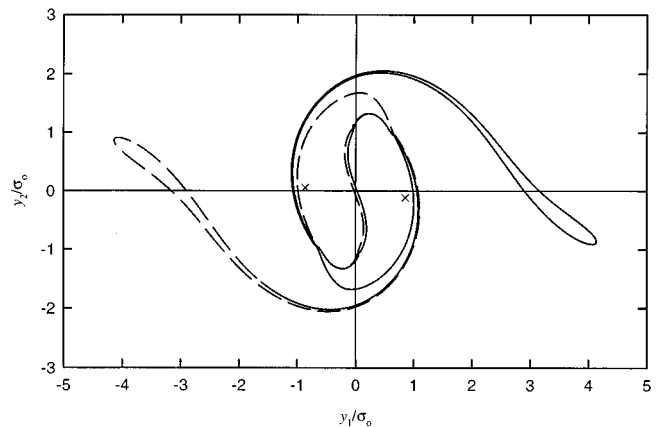
FIG. 14. Vortex core shapes during coalescence. $G/\sigma_0 = 3$, $J = 8.64$. ———: Initially trailing vortex; ––: initially leading vortex. $\times$: Vortex centroids.
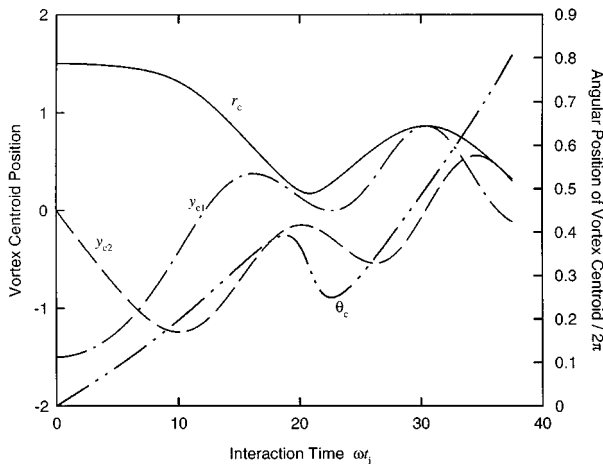
FIG. 15. Time variations of initially trailing vortex centroid positions. $G/\sigma_0 = 3$, $J = 8.64$. ———: r_c; —·—: θ_c; --: y_{c2}; —·—: y_{c1}.

FIG. 17. Time variations of initially trailing vortex accelerations. $G/\sigma_0 = 3$, $J = 8.64$. ———: $\ddot{y}_{c1}$; --: $\ddot{y}_{c2}$; —·—: angular acceleration $\ddot{\theta}_c$; —··—: radial acceleration $\ddot{r}_c$.

shown in Fig. 18. The rotating quadrupole nature of the sound field and its manner of rotation are similar to those in the previous case and are not discussed again. However, it is observed that the amplitude of the source strengths decreases at increased interaction time during which an overall decrease in the spatial separation between the vortex centroids is observed (Fig. 15), though there is an increase in the magnitudes of vortex centroid accelerations (Fig. 17). The far-field sound in this coalescence case does not seem to contain two components as in the previous leapfrogging case. Also, the magnitude of sound generated does not appear to be directly associated with the degree of vortex core deformation. Local minimum of source strength magnitude $\sqrt{\dddot{S}_1^2 + \dddot{S}_2^2}$ appears at $\omega t_i \approx 30$ while the maximum occurs at $\omega t_i \approx 9$ when the vortex core is less severely deformed (not shown here). This indicates that a higher degree of core deformation does not necessarily lead to greater sound radiation in the coalescence case. This is different from the results obtained in the leapfrogging case.

The smooth time variations of the source strengths $\dddot{S}_1$

and $\dddot{S}_2$ shown in Fig. 18 suggests that the vortex coalescence sound is not much related to the radial acceleration $\ddot{r}_c$ and the tangential acceleration $r_c\ddot{\theta}_c$ (Fig. 17). Their time variation frequencies coincide with those of the longitudinal and transverse velocities ($\dot{y}_{c1}$ and $\dot{y}_{c2}$) and accelerations ($\ddot{y}_{c1}$ and $\ddot{y}_{c2}$) (Figs. 16 and 17). The sound cannot come from the nonlinear couplings of these velocities or accelerations or both, as any such kind of couplings would increase the time variation frequency. Equation (16) is not applicable as the vortex centroids are not in circular motion in this case. The decrease in amplitudes of the source strengths $\dddot{S}_1$ and $\dddot{S}_2$ at increased interaction time suggests the importance of spatial separation in the sound generation mechanism as it is the only quantity that relates to vortex motion and shows a definite decreasing trend. The longitudinal and transverse velocities are not generally believed to be important in the generation of vortex sound, otherwise sound can be generated by steady vortex motion. However, they may affect the amplitude of the radiated sound. Thus it is expected that sound is generated by the longitudinal and transverse accelerating motions of the vortex centroids.

FIG. 16. Time variations of vortex centroid velocities. $G/\sigma_0 = 3$, $J = 8.64$. ———: $\dot{y}_{c1}$; --: $\dot{y}_{c2}$; —·—: angular velocity $\dot{\theta}_c$; —··—: radial velocity $\dot{r}_c$.
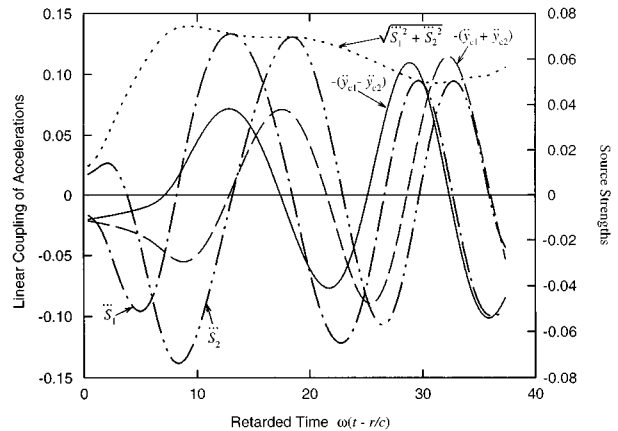
FIG. 18. Time variations of source strengths and linear couplings of accelerations. $G/\sigma_0 = 3$, $J = 8.64$. ———: $-(\ddot{y}_{c1} - \ddot{y}_{c2})$; --: $-(\ddot{y}_{c1} + \ddot{y}_{c2})$; —·—: $\dddot{S}_1$; —··—: $\dddot{S}_2$; ·····: $\sqrt{\dddot{S}_1^2 + \dddot{S}_2^2}$.
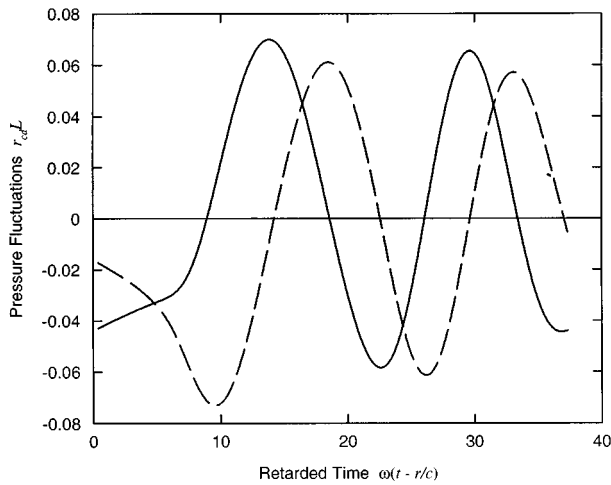
FIG. 19. Time variations of source terms and proposed approximations. $G/\sigma_0 = 3$, $J = 8.64$. ———: $-1.5r_{cd}(0.94\ddot{y}_{c1} - 0.34\ddot{y}_{c2})$; ——: $-1.5r_{cd}(0.34\ddot{y}_{c1} + 0.94\ddot{y}_{c2})$.

Results shown in Fig. 18 indicate that there exists phase shifts between the source and the longitudinal and transverse acceleration terms. The linear coupling of these accelerations is defined as $L = l_1\ddot{y}_{c1} + l_2\ddot{y}_{c2}$ where l_1 and l_2 are constants. Owing to the phase relationship between the acceleration terms (Fig. 17), these two constants represent a phase shift. Figure 18 illustrates the cases for $l_1 = -1$, $l_2 = 1$ and -1, indicating that source terms $\dddot{S}_i$ and L are nearly in phase. This further confirms the importance of the vortex centroid accelerations in the sound generation process. The time variation of $r_c L$ differs significantly from those of $\dddot{S}_1$ and $\dddot{S}_2$ and thus is not presented. However, the time variation pattern of vortex centroid radius r_c shown in Fig. 15 tends to suggest that r_c can be approximated by the sum of a roughly sinusoidal function r_{cs} and a slowly decreasing function r_{cd}. The amplitude of r_{cs} does not change much with time (not shown here) and thus, due to the decreasing nature of the sound amplitude, it is not believed to be important in the sound generation process. It is expected that r_c will become zero or very small as time approaches infinity and r_{cd} is assumed to take the form of $R_0 \exp(-\omega(t - r/c)/b)$. Here, R_0 denotes the initial radial distance of the vortex centroid ($= G/2\sigma_0$). However, the physical meaning of b is not definitely known. It may represent a certain kind of time scaling but it is more likely a modification factor of the r_c decrease rate. By least-squares method, b is estimated to be 29.2. The time variation of $r_{cd}L$ shows reasonable agreement with the source terms $\dddot{S}_1$ and $\dddot{S}_2$ (Figs. 18 and 19). To correct for the time shift, l_1 and l_2 are estimated to be -0.94 and 0.34, respectively for $\dddot{S}_1$. The corresponding values for $\dddot{S}_2$ are -0.34 and -0.94, respectively. The error for the fitting is around 20% for $\omega(t - r/c) > 10$. A similar situation is observed for other values of vorticity provided that coalescence occurs. It is not possible to relate the constants l_1 and l_2 to G/σ_0 as the spatial separation range for the occurrence of coalescence is very narrow. The two vortex cores touch each other at $G/\sigma_0 = 2$.

The results in this section indicate that the sound from vortex coalescence is mainly related to the longitudinal and transverse accelerating motions of the vortex centroids a short time after the commence of interaction (Figs. 18 and 19). Although an exact agreement in amplitude is not obtained, the fluctuating characteristics of the far-field sound can be explained by the linear coupling between the vortex centroid longitudinal and transverse accelerations.

III. CONCLUSIONS

The sound generated by the pairing of two inviscid finite-core-size equal two-dimensional vortices are studied numerically using the contour dynamics method and vortex sound theory. Two types of pairing, namely the leapfrogging motion and coalescence, are discussed. The sound generation mechanism for each type of pairing is examined in terms of the motions of the vortex centroids as these motions are more easily studied experimentally than those suggested in existing literature.

The sound fields resulted from leapfrogging motion and coalescence are lateral quadrupoles. However, they rotate with a speed different from that of the vortex field. Their directions of rotation follow that of the vortex field for the majority of time. This is different from the results in existing literature on thin core vortex pairing sound study, which indicate that the vortex and the sound fields rotate in the same direction.

In the leapfrogging case, the vortex centroids move in a circular orbit without serious core deformation. The far-field sound is found to consist of two major components—one of low frequency and the other of high frequency. The low-frequency component is generated by the nonlinear couplings of the longitudinal and transverse accelerations of the vortex centroids. The high-frequency component is related to the coupling of the longitudinal and transverse velocities with the radial and angular accelerations of the vortex centroids. The larger the initial separation of the vortex centroids, the more important this high-frequency component becomes. Therefore, the accelerating and decelerating motions of the vortices play an important role in the generation of sound in the leapfrogging case.

The sound generated by vortex coalescence does not contain two components of different frequency as in the leapfrogging case. In general, the spatial separation between the vortex centroids decreases gradually with time. Results suggest that the sound so generated is associated with the longitudinal and transverse accelerations of the vortex centroids and there being no evidence that radial and angular accelerations of these centroids are related to the sound generation process. While linear couplings of vortex centroid longitudinal and transverse accelerations do not give the exact amplitudes of the sound source terms, they show reasonable agreement with the fluctuating characteristics of the sound source strength. The slowly decreasing component of the spatial separation between vortex centroids is also found to be important in controlling the amplitude of the far-field sound.

[1] M. J. Lighthill, ''On sound generated aerodynamically I. General theory,'' Proc. R. Soc. London, Ser. A **211**, 564–587 (1952).

[2] N. Curle, ''The influence of solid boundary upon aerodynamic noise,'' Proc. R. Soc. London, Ser. A **231**, 505–514 (1955).

[3] A. Powell, ''Vortex sound theory,'' J. Acoust. Soc. Am. **36**, 177–195 (1964).

[4] M. S. Howe, ''Contributions to the theory of aerodynamic sound, with application to excess jet noise and the theory of the flute,'' J. Fluid Mech. **71**, 625–673 (1975).

[5] J. E. Bridges and A. K. M. F. Hussain, ''Roles of initial conditions and vortex pairing in jet noise,'' J. Sound Vib. **117**, 289–311 (1987).

[6] M. H. Krane and W. P. Pauley, ''Estimation of the direct acoustic radiation from a transitional boundary layer using velocity measurements,'' J. Sound Vib. **181**, 737–763 (1995).

[7] F. Obermeier and K. Q. Zhu, ''Sound generation by rotor-vortex interaction in low Mach number flow,'' J. Aircraft **30**, 81–87 (1993).

[8] P. A. Lush, ''Measurements of subsonic jet noise and comparison with theory,'' J. Fluid Mech. **46**, 477–500 (1971).

[9] D. G. Crighton, ''Basic principles of aerodynamic noise generation,'' Prog. Aerosp. Sci. **16**, 31–96 (1975).

[10] W. Möhring, ''On vortex sound at low Mach number,'' J. Fluid Mech. **85**, 685–691 (1978).

[11] K. Shariff, A. Leonard, and J. H. Ferziger, ''Dynamics of a class of vortex rings,'' NASA Tech. Memo. **TM-102257** (1989).

[12] S. K. Tang and N. W. M. Ko, ''On sound generated by the interaction of two inviscid vortex rings moving in the same direction,'' J. Sound Vib. **187**, 287–310 (1996).

[13] D. J. Lee and S. O. Koo, ''Numerical study of sound generation due to a spinning vortex pair,'' AIAA J. **33**, 20–26 (1995).

[14] B. E. Mitchell, S. K. Lele, and P. Moin, ''Direct computation of the sound from a compressible corotating vortex pair,'' J. Fluid Mech. **285**, 181–202 (1995).

[15] N. J. Zabusky, M. H. Hughes, and K. V. Roberts, ''Contour dynamics for the Euler equations in two dimensions,'' J. Comput. Phys. **30**, 96–106 (1979).

[16] D. G. Dritschel, ''The nonlinear evolution of rotating configurations of uniform vorticity,'' J. Fluid Mech. **172**, 157–182 (1986).

[17] C. Pozrikidis and J. J. L. Higdon, ''Nonlinear Kelvin-Helmholtz instability of a finite vortex layer,'' J. Fluid Mech. **157**, 225–263 (1985).

[18] J. E. Ffowcs Williams and D. L. Hawkings, ''Shallow water wave generation by unsteady flow,'' J. Fluid Mech. **31**, 779–788 (1968).

[19] H. Ziegler, *Mechanics Vol. 1 Statics of Rigid Bodies, Fluids, and Deformable Solids* (Addison-Wesley, Reading, MA, 1965).

[20] H. Lamb, *Hydrodynamics* (Dover, New York, 1932).

[21] D. G. Dritschel, ''A general theory for two-dimensional vortex interactions,'' J. Fluid Mech. **293**, 269–303 (1995).

Short pulse acoustic excitation of microbubbles

Nicholas G. Pace, Alastair Cowley,[a] and A. M. Campbell[b]

School of Physics, University of Bath, Bath BA2 7AY, United Kingdom

When a bubble is excited at or near its natural frequency it is a very effective scatterer and absorber of sound. In fact at resonance, scattering and absorption cross sections are of the order of 1000 times greater than the geometric cross section for typical marine microbubbles. As with any resonant system, a bubble takes a finite, albeit short time to ring up to steady-state oscillation and continues to oscillate for a finite time after the driving force ceases. A number of studies have been carried out to attempt to describe how acoustic absorption and backscatter depend on the duration of the driving force. Most notable of these are the backscatter measurements for short pulses in near surface sea water [Akulichev *et al.*, Sov. Phys. Acoust. **23**(3), 177–180 (1986)] and the forward transmission laboratory measurements for short pulse lengths at 120 kHz [H. R. Suiter, J. Acoust. Soc. Am. **91**, 1383–1387 (1992)]. Akulichev observed a decrease in acoustic backscatter for short pulses at three discrete frequencies, whereas Suiter found no corresponding enhancement in forward transmission. This paper details a comprehensive study which measured acoustic attenuation through a well defined bubble cloud over the frequency range 20–200 kHz and for pulse lengths from 20 cycles down to a single cycle using a parametric transmitter. The experiment simulated the conditions for which a decrease in attenuation with decreasing pulse length might have been expected. No effect was observed for two different but well-defined bubble distributions. © *1997 Acoustical Society of America.* [S0001-4966(97)04207-0]

PACS numbers: 43.30.Gv [JHM]

INTRODUCTION

When a bubble is near or at its natural frequency it is a very effective scatterer and absorber of sound. In fact at resonance, scattering and absorption cross sections are of the order of 1000 times greater than the geometric cross section for typical marine microbubbles. As with any resonant system, a bubble takes a finite time to achieve a steady-state oscillation: It also continues to oscillate for a finite time after the driving force ceases. There are some studies of the dependence of the acoustic absorption and backscatter on the duration of the driving forces. Akulichev[1] studied acoustic backscatter from bubbles in near-surface sea water as a function of the incident acoustic pulse length. Suiter[2] measured the forward transmission through bubble clouds in the laboratory for relatively short pulse lengths. Akulichev[1] observed a decrease in the acoustic backscatter at 5, 15, and 35 kHz, whereas Suiter[2] found no enhancement or pulse length related change in the forward transmission at 120 kHz.

In this paper an account of a comprehensive laboratory study of the acoustic transmission loss through bubble clouds is presented. A parametric source[3] was employed which allowed measurements to be made over the frequency range 20 kHz–200 kHz with pulse lengths from 20 cycles down to 1 cycle. The experiment embraced conditions for which a decrease in attenuation with decreasing pulse length might have been expected. However, no such effect was observed.

I. OUTLINE THEORETICAL BASIS

The steady-state extinction cross section[4] of a single bubble of radius a is

$$\sigma_e = \frac{4\pi a^2 (\delta/ka)}{((f_0/f)^2 - 1)^2 + \delta^2},\tag{1}$$

where f is the sound frequency, k is the wave number, and δ is the damping factor. The resonance frequency f_0 is approximately $3.25/a$ Hz at shallow depths, where a is in meters. The attenuation α in dB per meter is

$$\alpha(f,t) = 4.34 \int_0^\infty \{\sigma_{nr}(f) + \sigma_r(f,t)\} \psi(a)\,da,\tag{2}$$
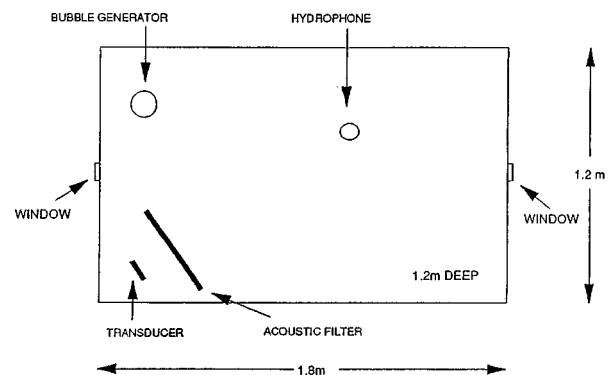
FIG. 1. The experimental arrangement for the measurement of acoustic transmission loss through bubble clouds. The windows indicated are for the optical measurement of the bubble size distribution via a laser based diffraction technique not reported here.
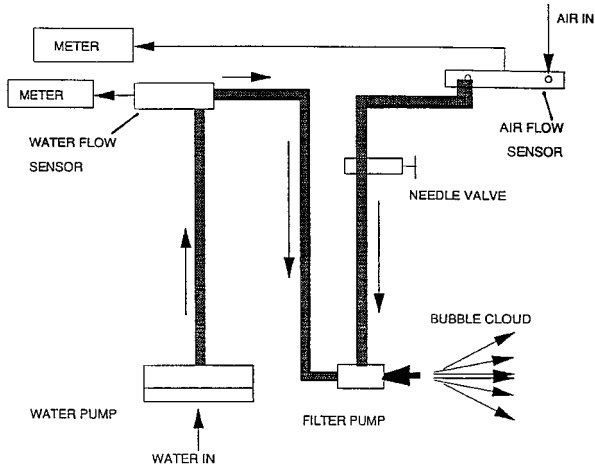
FIG. 2. Arrangement for the generation of the bubble clouds using a filter pump. The air flow and water flow into the pump are carefully monitored and reproducible.

where $\psi(a)$ is the bubble radius distribution function and σ_{nr} and σ_r are the nonresonant and resonant components of the extinction cross section, respectively. For very short pulses the resonant contribution is expected to be significant and the time-dependent effects to be noticeable.

The time dependence of the extinction cross section may be written as

$$\sigma_e(t) = \sigma_{nr} + (1 - \exp(-2t/t_0))\sigma_r, \qquad (3)$$

in which t_0 is the time required for the bubble to execute
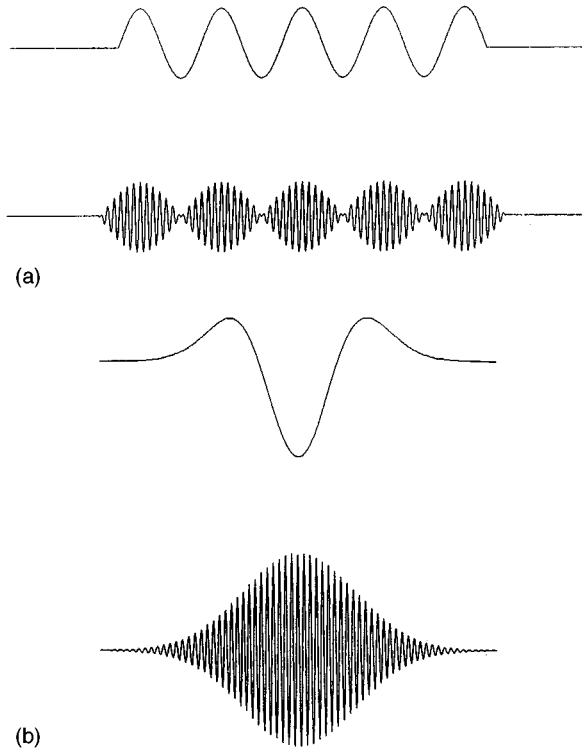
FIG. 3. The application of a modulated high-frequency signal to the acoustic transducer results in a low-frequency signal in the water due to nonlinear processes in the water. (a) Shows the production of a long pulse and (b) shows the production of a wide bandwidth pulse.
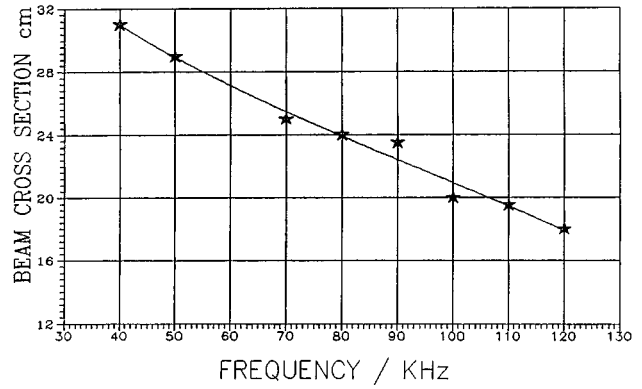
FIG. 4. The characteristics of the acoustic field produced via the parametric array terminated by the acoustic filter shown in Fig. 1 were carefully measured. This figure shows the -3-dB cross-sectional width of the difference frequency beam as a function of frequency. It was measured at 40 cm from the acoustic filter when the transmitting transducer to filter distance was 40 cm. (The points are measurements and the line is a least squares fit to the measured points.)

Q/π oscillations, Q being the quality factor for the bubble resonance. The experiments reported below utilize pulse lengths small compared with t_0, and thus if time-dependent effects for the bubble cloud as a whole occur they should be observable.

II. EXPERIMENTAL ARRANGEMENT AND PROCESSING

The experiments were performed in a laboratory tank 1.2 m×1.8 m×1.2 m deep, as shown in Fig. 1. The transmitting acoustic transducer was located in one corner and directed through a region in which bubbles were generated toward a small hydrophone (B&K 8103). The geometrical arrangement and the repetition rate of the acoustic transmis-
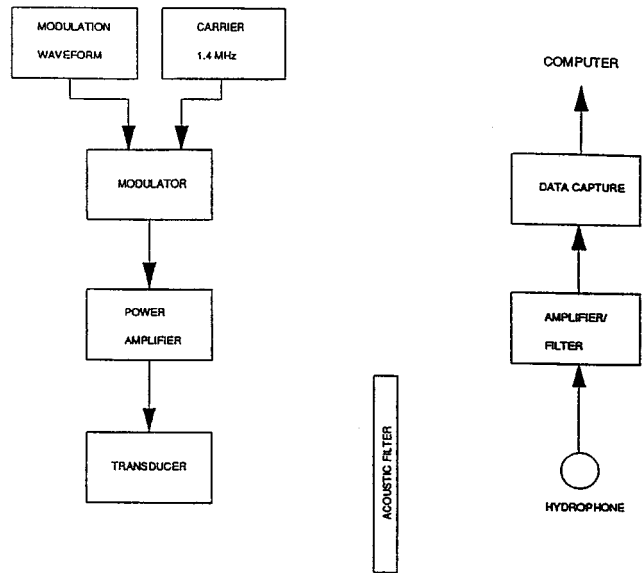
FIG. 5. Block diagram of the arrangement for transmitting and receiving the acoustic signals. The modulation waveform was read out from a storage device and was designed to produce a long sequence of difference frequency signals as detailed in Fig. 6 and the text.

LONG   SHORT   LONG   SHORT

Freq   F1   F1   F2   F2

No OF CYCLES   N   n   N   n
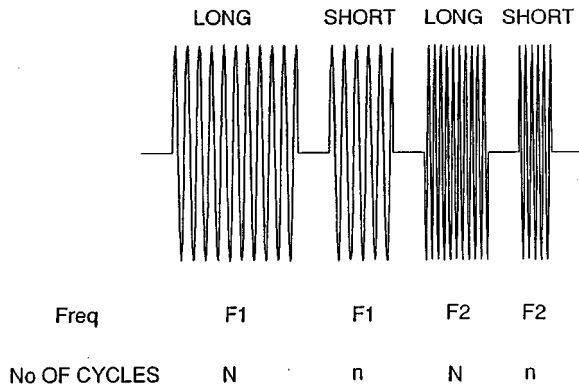
FIG. 6. An example of the sequence of difference signals produced by the stored modulation. Long pulses (usually 20 cycles) are successively transmitted at frequency intervals of 1 kHz interleaved by shorter pulses at the same frequencies. In this way the frequency range from 20 to 200 kHz was covered. The minimum time interval between pulses was set by the tank size to be 20 ms. The long pulse acted as a continuously present reference calibration.

sions were chosen carefully to avoid reflections from the tank walls occurring in the time gate of interest.

The bubble generation was accomplished in an approximately repeatable manner by passing air and water through an orifice in what is known as a filter pump. The air and water flows were monitored as indicated in Fig. 2. The bubble cloud developed to occupy the tank, but was excluded by means of thin plastic sheets from the region of parametric generation.

Provision was made for a simultaneous optical measurement of the bubble size distribution using a laser based diffraction technique. However, the method was only partially successful. While confirmation of the shape of the bubble size distribution was obtained, absolute values were unreliable and no results are presented.

The broadband width and short pulse capability required of the acoustic transmitter was achieved by employing a parametric source.[3] An air-backed 80-mm-diam piezoelectric ceramic disk with a thickness mode resonance of 1.4 MHz was excited by a modulated waveform as depicted in Fig. 3(a). The nonlinear processes which occur in the water result
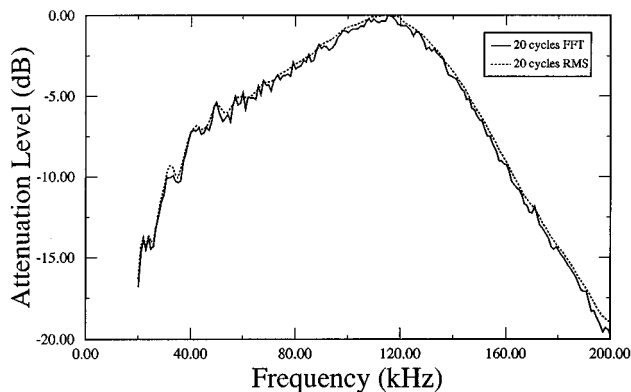
FIG. 7. The frequency response of the system is shown, simply obtained by running the experiment in the absence of bubbles. The curves shown were obtained using 20 cycle pulses and analyzed both by the rms and the FFT approach, as described in the text. The same curves are obtained by frequency analysis of broadband pulses [see Fig. 3(b)].
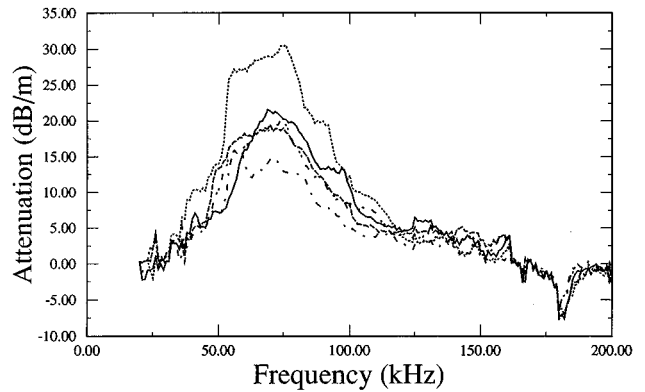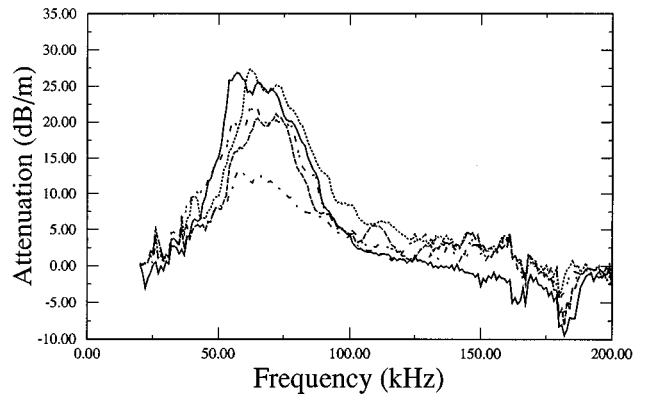
FIG. 8. These two figures each show five consecutive measurements of the attenuation spectra through the same nominal bubble cloud obtained using the "long" 20 cycle pulses. The time to acquire each spectrum was about 3.5 s while the minimum time between acquiring spectra was 40 s.
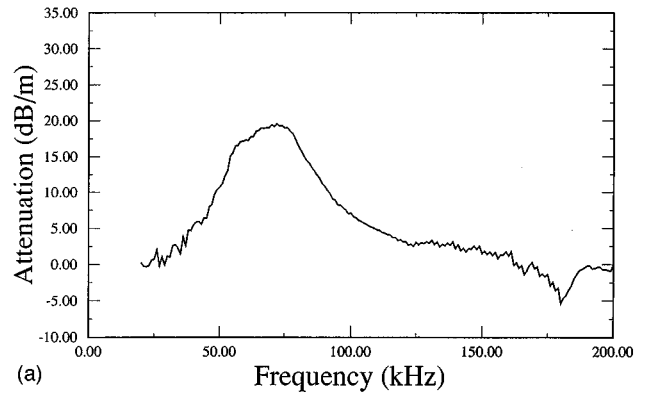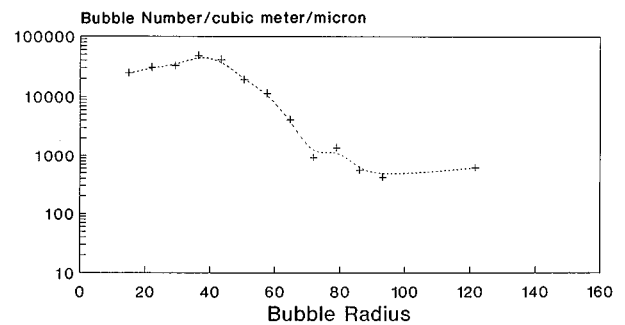
FIG. 9. (a) The average of 20 attenuation spectra required using the "long" 20 cycle pulses is shown. (b) The inversion of the spectrum of (a) performed at the indicated points. The solid line is simply a smooth curve through the inversion points to aid the interpretation.
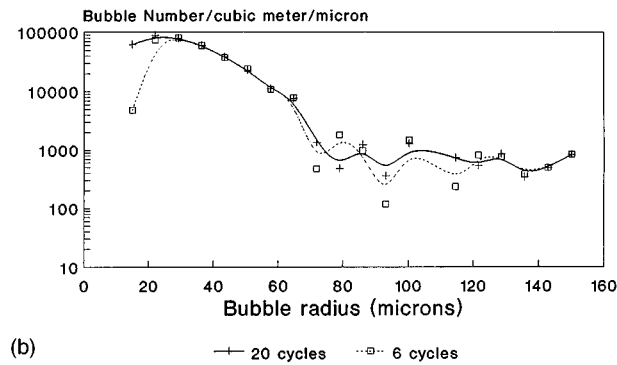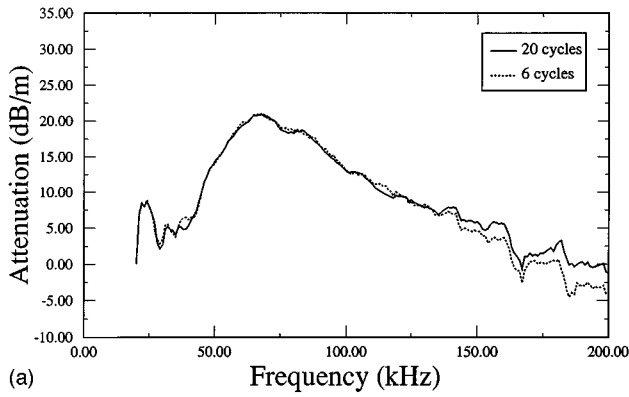
FIG. 10. (a) The average of 16 attenuation spectra acquired using interleaved "long" 20 cycle pulses and "short" 6 cycle pulses is shown. (b) The inversion of the spectrum of (a) performed at the indicated points. The solid line is simply a smooth curve through the inversion points to aid the interpretation.
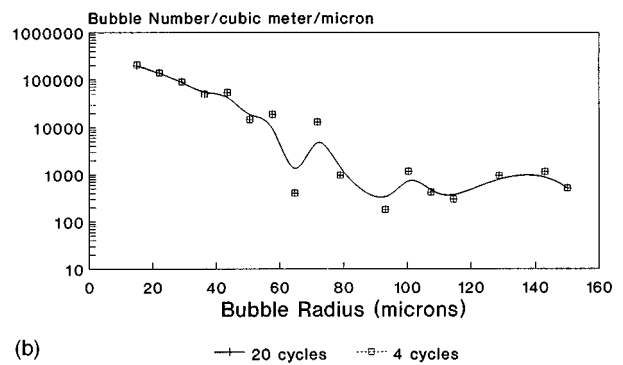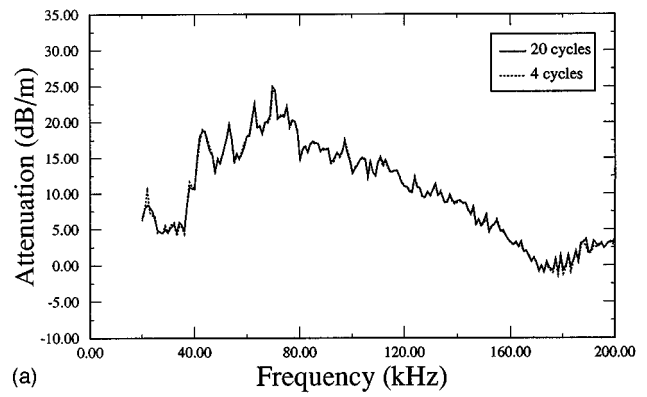
FIG. 11. (a) Single attenuation spectra acquired using interleaved "long" 20 cycle pulses and "short" 4 cycle pulses is shown. (b) The inversion of the spectrum of (a) performed at the indicated points. The solid line is simply a smooth curve through the inversion points to aid the interpretation.

in the generation of a low-frequency signal at twice the frequency of the modulation. This low-frequency signal is generated throughout the volume occupied by the carrier wave, phased in the manner of an endfire array. As the number of cycles in the carrier wave modulation is decreased, the low-frequency signal approaches an approximation to a single cycle, as shown in Fig. 3(b). The situation depicted in Fig. 3(b) is sometimes known as self-demodulation.[3] The generation of the low-frequency signal is restricted in the experiment by inserting an acoustic filter as shown in Fig. 1. This absorbs the 1.4-MHz carrier signal while allowing the low-frequency signal to pass through. This manner of employing a parametric array as a convenient laboratory tool wherein the generation region is separated from the test region by use of acoustic filters is well established and is reviewed in Ref. 3.

The spatial characteristics of the difference frequency beam in the test region were measured. Some results are summarized in Fig. 4 where the beam cross-sectional width (-3 dB) is plotted against difference frequency.

It was necessary to run the transmission loss experiments for one particular pulse length over the complete range of frequencies during a time interval in which the evolution of the bubble cloud was slight. This was accomplished by storing the modulation waveform sequence in a memory and applying this to the modulator and hence transducer (Fig. 5) so that the acoustic pulse train shown in Fig. 6 was produced in the water. This consists of an interleaving of a "long"

pulse with a "short" pulse at the same frequency and then repeating at a different frequency. In this way the frequency range from 20 to 200 kHz was covered in 1-kHz steps during one pulse sequence. The "long" pulse was always 20 cycles while the "short" pulse although a fixed number of cycles for one transmission sequence, varied from 20 cycles down to 1 cycle. The "long" pulse served as a reference. The normalized system frequency response was obtained by transmitting a pulse sequence in the absence of any bubbles. This is shown in Fig. 7.

Measurements made in the presence of the bubble cloud were converted to transmission loss using the system response shown in Fig. 7. Attenuation values were obtained from transmission loss by dividing by the acoustic filter/receiver distance. After traversing the bubble cloud, the acoustic pulse sequence was received on the hydrophone and passed to an analogue to digital converter and stored for processing in a computer. The time taken to transfer the digitized pulse sequence to the computer was such that the time between captured pulse sequences was about 40 s. However, the time duration of a pulse sequence was only about 3.5 s.

When a pulse is several cycles long a meaningful rms amplitude can be calculated. As the number of cycles in a pulse is reduced toward one, an appropriate measure of its amplitude is obtained by calculating the amplitude of the Fourier component at its center frequency using a rectangular window. The latter is defined so that it coincides with the rms amplitude when the pulse contains six or more cycles.

1477   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Pace *et al.*: Excitation of bubbles   1477
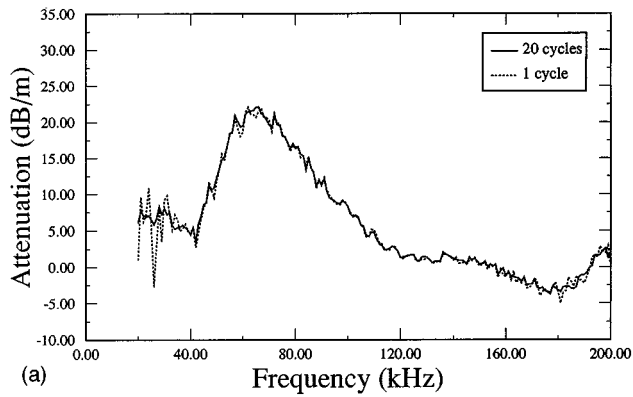
(a)
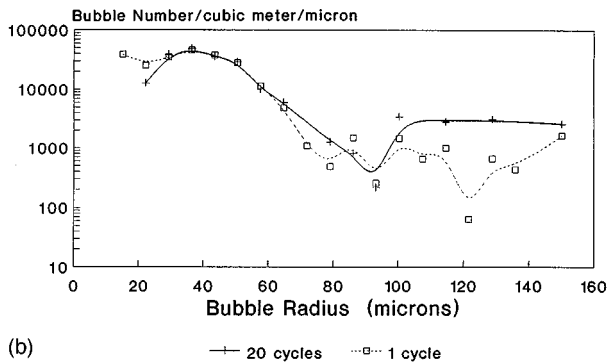
(b)

FIG. 12. (a) The average of 10 attenuation spectra acquired using interleaved ''long'' 20 cycle pulses and ''short'' 1 cycle pulses is shown. (b) The inversion of the spectrum of (a) performed at the indicated points. The lines are simply smooth curves through the inversion points to aid the interpretation.

For pulses of appropriate lengths, pulse sequences processed by either method gave the same results.

III. RESULTS

Figure 8(a) and (b) shows the attenuation through an evolving bubble cloud for 20 cycle pulses. Each of the graphs in Fig. 8 shows five sequences. As mentioned above, the time during a pulse sequence is only 3.5 s while the time between sequences is 40 s. The average for 20 pulse sequences for the ''same'' bubble cloud is shown in Fig. 9(a). Figure 9(b) shows a full inversion of the attenuation spectra presented in Fig. 9(a) in terms of bubble number as a function of size. The basis for the inversion is first to smooth the attenuation spectrum to reduce the experimental fluctuations, then using the Fredholm inversion technique[5] to use a single value decomposition method to derive the bubble number over the size range 15–150 μm radius. This technique proposed by Moritz and Commander[5] uses a sophisticated choice of basis functions to describe the bubble radial distribution function. This approach can be reliably applied up to approximately 150 kHz.

Figures 10, 11, and 12 show that the frequency dependence of the attenuation through bubble clouds as the pulse length is reduced from 6 to 4 to 1 cycle. The interleaving of ''short'' and ''long'' pulses in the pulse sequences allows the ''long'' pulse 20 cycle results to be obtained simultaneously with the ''short'' pulse results. It is not always possible to distinguish between the ''long'' and ''short'' pulse
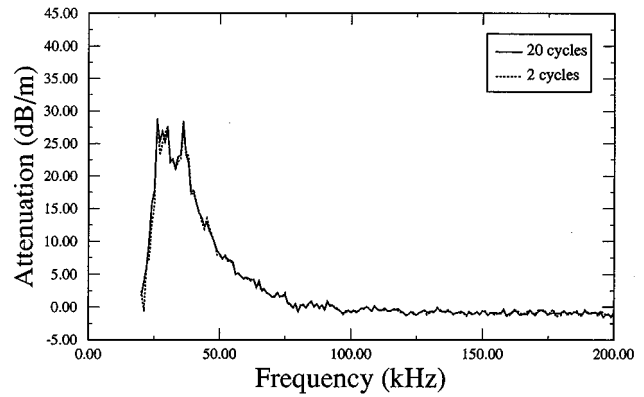
FIG. 13. The average of 10 attenuation spectra acquired using interleaved ''long'' 20 cycle pulses and ''short'' 2 cycle pulses is shown when the air flow and water flow to the filter pump has been adjusted to significantly reduce the number of bubbles in the smaller size range compared with the clouds of Figs. 8–12.

results. The inversion to bubble size distribution is given in all cases. Since the impulse response time for a resonant bubble is likely to be greater at lower frequencies, corresponding to larger bubble size, the water and air flow rates were adjusted to significantly reduce the number of bubbles in the smaller size range.

Figures 13 and 14 present attenuation spectra for ''long'' 20 cycle pulses and ''short'' 2 and 1 cycle pulses, respectively, through this larger bubble size bubble cloud. Inversion of the spectrum of Fig. 13 is presented in Fig. 15. The much reduced number of small bubbles is clearly shown during the measurement sequence. Again, no difference between the ''short'' pulse results and the ''long'' 20 cycle reference pulse results are seen.

IV. DISCUSSION AND CONCLUSIONS

Theoretical considerations and practical requirements initially lead us to believe that due to the finite, albeit short time required for a bubble to ''ring up'' to steady-state resonant behavior, a reduced attenuation for a short pulse transmitted through a bubbly medium might occur. To investigate this possibility, a reproducible bubble cloud similar to that created by breaking waves or ships wakes was generated in
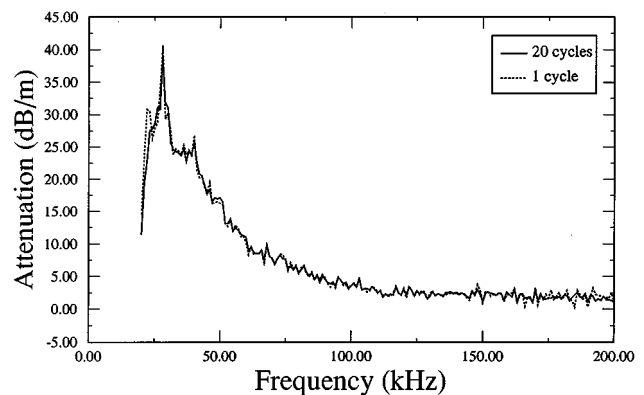
FIG. 14. The average of 10 attenuation spectra acquired using interleaved ''long'' 20 cycle pulses and ''short'' 1 cycle pulses is, shown for the same nominal bubble cloud as for Fig. 13.
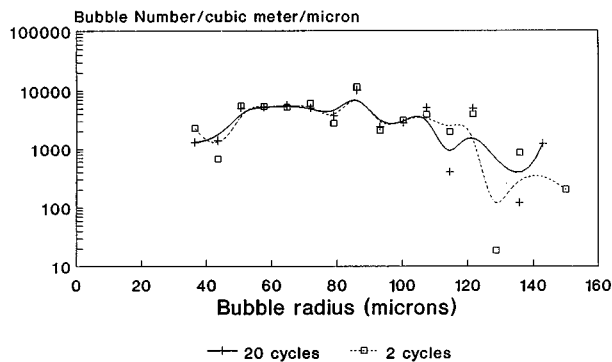
FIG. 15. The inversion of the spectrum of Fig. 13 performed at the indicated points. The lines are simply smooth curves through the inversion points to aid the interpretation.

controlled laboratory conditions. The acoustic attenuation spectra were measured over the wide frequency range of 20–200 kHz, for both ''long'' pulses of 20 cycles down to ''short'' pulses of a single cycle, the whole experiment being concluded on a time scale shorter than that of the bubble cloud evolution. For two different bubble distributions no significant differences were observed between the ''long'' and ''short'' pulse cases.

It is concluded that acoustically exciting a wide range of bubble sizes rather than a single bubble size may mask the transient effects. Furthermore, due to multiple scattering, each individual bubble is excited not only by the incident pressure field, but also by the sum of the field scattered by its near neighbors. This loss of coherence may also mask the effect of transient behavior. Thus acoustic systems transmitting a series of very short pulses or frequency sweeps are unlikely to result in significantly lower attenuation through a bubble cloud than a conventional long pulse system.

ACKNOWLEDGMENT

[1] V. A. Akulichev, V. A. Bulanov, and S. A. Klenin, ''Acoustic sensing of gas bubbles in the ocean medium,'' Sov. Phys. Acoust. **32**, 177–180 (1986).

[2] H. R. Suiter, ''Pulse length effects on the transmissivity of bubbly water,'' J. Acoust. Soc. Am. **91**, 1383–1387 (1992).

[3] V. F. Humphrey, ''Nonlinear acoustics as a laboratory tool,'' Proc. Inst. Acoust. **14**, 99–113 (1992).

[4] C. S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1977).

[5] E. Moritz and K. W. Commander, ''Solving the inverse scattering problem to obtain acoustic bubble spectra,'' J. Acoust. Soc. Am. Suppl. 1 **84**, S186 (1988).

Choosing Biot parameters for modeling water-saturated sand

C. J. Hickey and J. M. Sabatier

National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

The work of Chotiros [J. Acoust. Soc. Am. **97**, 199–214 (1995)] has stimulated renewed interest in detecting the Biot type II (slow) *P* wave for the purpose of ocean bottom characterization. Chotiros postulates that total internal reflection of the type I (fast) *P* wave can occur at a water–sediment interface and that the observed refracted energy below the critical angle is associated with the type II (slow) *P* wave. The choice of parameters used in the Biot theory to model these observations is discussed. Two major differences between the parameter set proposed by Chotiros and previous parameter sets for sandy sediments exist. First, the value for the unjacketed bulk modulus, K_r, is about five times less than the bulk modulus for quartz crystals, which is commonly used. This value is obtained by equating it to a value of effective grain modulus measured by Molis and Chotiros [J. Acoust. Soc. Am. **91**, 2483(A) (1992)]. Second, the value for the frame bulk modulus, K_b, obtained by fitting to measured values of the type I (fast) *P* wave speed, is from 50 to 100 times larger than in the previous models. The larger value for the frame bulk modulus is in part a consequence of the smaller value of the unjacketed bulk modulus. Within the constraints imposed by the Biot theory, subject to the assumption of an effective homogeneous solid, the proposed values of unjacketed bulk modulus and frame bulk modulus are incompatible. It is shown here that it is not possible to predict Chotiros' observed large phase speed of the type II *P* wave by assuming an equivalent homogeneous solid. However, it is possible to choose a compatible set of parameters for the Biot theory so as to fit the velocity observations once the assumption of an equivalent homogeneous solid is relaxed. © *1997 Acoustical Society of America.* [S0001-4966(97)06309-1]

PACS numbers: 43.30.Ma, 43.30.Zk [JHM]

I. INTERPRETATION OF BIOT'S EQUATIONS

Biot[1,2] proposed a set of coupled differential equations to describe wave propagation in fluid-filled porous media. The ''elastic parameters'' introduced in this work were not clearly defined in terms of measurable quantities. Biot and Willis[3] published a later paper giving an interpretation of these ''elastic parameters'' in terms of measurable quantities and placed constraints on the possible values of some of these ''elastic parameters.'' Over the years the equations proposed by Biot have been written in many forms. Biot's equations for dilatational waves used by Chotiros[4] are

$$\mu \nabla^2 \mathbf{u} + (H-\mu)\nabla(\nabla \cdot \mathbf{u}) \cdot C\nabla(\nabla \cdot \mathbf{w}) = \rho \ddot{\mathbf{u}} - \rho_f \ddot{\mathbf{w}} \qquad (1)$$

and

$$C\nabla(\nabla \cdot \mathbf{u}) - M\nabla(\nabla \cdot \mathbf{w}) = \rho_f \ddot{\mathbf{u}} - (c\rho_f/\beta)\ddot{\mathbf{w}} - (F\eta/\kappa)\dot{\mathbf{w}}. \qquad (2)$$

Starting with the original equations of Biot[1,2] and the expressions presented by Biot and Willis,[3] the elastic parameters are written in terms of measurable moduli as

$$M = \frac{\mathbf{K}_r}{1 + K_r/K_r^f - K_b/K_r}, \qquad (3a)$$

$$C = \alpha M, \qquad (3b)$$

$$H = \alpha C + K_b + \tfrac{4}{3}\mu, \qquad (3c)$$

$$\alpha = 1 - K_b/K_r, \qquad (3d)$$

where μ is the shear modulus of the frame, K_b is the bulk modulus of the frame, K_r is a bulk modulus measured in an unjacketed bulk experiment (believed to be the bulk modulus of the grains), and K_r^f is the inverse of the coefficient of fluid content as defined by Biot and Willis.[3] Biot and Willis[3] also argue for the following constraint:

$$\beta < \alpha < 1, \qquad (4)$$

where β is the porosity. In order to understand the origins of this constraint better, let's substitute Eq. (3d) into Eq. (4) and obtain the possible range of values for the bulk modulus of the frame as

$$0 < K_b < (1-\beta)K_r, \qquad (5)$$

Now consider a quasistatic compressional deformation of the type used to measure frame bulk modulus. Essentially, the sample is placed in an impermeable but deformable jacket, and changes in the volume of the sample are measured as the pressure is increased on the jacket. The pressure is usually exerted by some external confining fluid and is therefore referred to as confining pressure. The pore fluid pressure is kept constant. The lower limit on the frame bulk modulus, K_b, i.e., zero bulk modulus, would be for an idealized sample which has no shear strength. When an external confining pressure is imposed on such an idealized sample, its solid framework collapses through shear deformation. Therefore, a large volume change is obtained with little applied pressure, yielding, in the limit, a zero value for the bulk modulus. The upper limit on frame bulk modulus corresponds to an idealized sample which has infinite shear strength. In this case the external confining pressure cannot cause any shear deformation of the grains or frame because

of the infinite shear strength. Therefore, the sample must undergo pure volumetric deformation. Since the pore fluid does not sustain any of the external load, all the deformation is within the solid frame, which is simply related to the unjacketed bulk modulus, K_r. The factor of $1-\beta$ comes about due to the assumption that the solid component makes up that fraction of the total sample surface area on which the external confining pressure is applied.

The expressions for the "parameters" in Biot's equations given by the equations (3a), (3b), (3c), and (3d) above are different from Chotiros' Eqs. (3), (4), and (5). A relationship for the coefficient of fluid content is required in order to make the two sets of expressions the same. This relationship [Biot and Willis,[3] Eq. (24)] is

$$\frac{1}{K_r^f} = \beta \left(\frac{1}{K_f} - \frac{1}{K_r} \right).$$

(6)

The necessary condition for this relationship to hold is that the sample undergo no change in porosity during an unjacketed compression. Biot and Willis[3] state that the above relation (6) is true "for the special case in which the material of the porous matrix is homogeneous and isotropic and the fluid completely saturates the pores." For such a special case Nur and Byerlee[5] have shown that

$$K_r = K_s,$$

(7)

where K_s is the bulk modulus of the solid grains. Several authors[6-11] have noted that a relationship requiring the unjacketed compressibility of the frame be equal to the compressibility of the grains is very restrictive.

Biot and Willis[3] argue that a rock composed of heterogeneous matrix material may in effect behave as some equivalent homogeneous rock. More specifically, the heterogeneous matrix undergoes the same strain as the pores in an unjacketed test. Berryman[11] demonstrated that effective-stress data on fluid transport through porous rocks cannot be explained in terms of such an equivalent homogeneous rock. However, let's examine further the possibility of an equivalent homogeneous rock.

For a rock with heterogeneous matrix material it should be possible to place bounds on the effective grain bulk modulus using the Voigt–Reuss–Hill[12] formulae. Under the assumption of an equivalent homogeneous solid these are also bounds on the unjacketed bulk modulus. The range is given by

$$K_{\text{Reuss}} \leq K_{\text{eff}}^s \leq K_{\text{Voigt}},$$

(8)

where

$$\frac{1}{K_{\text{Reuss}}} = \sum_{i=1}^n \frac{c_i}{K_i}$$

(9)

and

$$K_{\text{Voigt}} = \sum_{i=1}^n c_i K_i,$$

(10)

n being the number of different solids, K_i the bulk modulus of a particular solid component, and c_i is the volume fraction of the solid component with respect to the total volume of

solid and not the sample volume. This type of volume fraction must be used because we are estimating the effective grain bulk modulus and not the undrained bulk modulus, in which case the pore fluid would have to be taken into account. The Reuss limit is based on the fact that all the different grains support the same stress, whereas the Voigt limit assumes that all the different grains undergo equal deformation.

For a material to behave as the equivalent homogeneous solid as required by Biot and Willis,[3] the heterogeneous matrix must undergo the same strain as the pores in an unjacketed test. Having a matrix undergo the same strain as the pores is true of the deformation in the Voigt limit, but the required applied stress is not typical of an unjacketed test. In an unjacketed test each grain supports the same pressure, as is usually argued, since the confining pressure and pore pressure are equal, so that the unjacketed moduli for an equivalent homogeneous sample should be equal to that of the Reuss limit. During an unjacketed compression of a porous matrix, however, relative compression of grains and the inability for these grains to slip is likely to produce deviatoric strains within the grains. This will cause changes in the shape of the grains, and therefore the strain of the matrix will not, in general, be equal to the strain of the pores. However, the point of Biot and Willis[3] is that there might exist a special case where changes in shape due to deviatoric strains "cancel out" and the matrix strain equals the pore strain. For such a case the value of the unjacketed bulk modulus should lie with the Voigt–Reuss–Hill bounds.

II. EXAMINATION OF CHOTIROS' VALUES FOR MODULI

The majority of the modeling using the Biot theory assumes an effective homogeneous solid so that the coefficient of fluid content can be calculated via Eq. (6) and the equality between unjacketed and grain bulk moduli, Eq. (7), can be used. Furthermore, the grain bulk modulus commonly used is the solid bulk modulus, K_s, of the prominent mineral. For the majority of sands, the bulk modulus of a quartz crystal, which has a value around 3.6×10^{10} Pa, is used. The value of unjacketed bulk modulus, K_r, used by Chotiros[4] is 7×10^9 Pa. This value was based on a measurement of an effective grain modulus by Molis and Chotiros.[13] Given the composition of samples (Chotiros,[4] Table II), the most obvious reason for the different values of grain bulk modulus is the inhomogeneity of the solid matrix.

From the mineral composition of the sand samples the range of effective grain bulk modulus for the ARL:UT sand tank can be calculated based on 86% quartz ($K_{\text{qrtz}} = 3.6 \times 10^{10}$ Pa), 13% feldspar ($K_{\text{feld}} = 4.8 \times 10^{10}$ Pa) and 1% other, say with a bulk modulus equal to that of water ($K_w = 2.25 \times 10^9$ Pa). The 1% could be considered isolated water-filled pores. This might seem unrealistic but it will push the lower bound of the range of bulk moduli to a smaller extreme. The range of bulk moduli for the ARL:UT sand tank is

$$3.3 \times 10^{10} \text{ Pa} \leq K_s^{\text{ARL:UT}} \leq 3.7 \times 10^{10} \text{ Pa}.$$

(11)

The predicted range in effective grain bulk modulus for the sample is quite small and the value measured by Molis and Chotiros[13] lies below the lower bound. One scenario for which the value of the effective grain bulk modulus would be within predicted bounds is to include a small fraction of a very compressible component such as air ($K_{air} = 1.5 \times 10^5$ Pa). This is somewhat far fetched because it requires that this air component constitute part of the solid framework and be distributed such that it can support equal amounts of stress. The above discussion raises questions about the measurement of Molis and Chotiros.[13] A more detailed description of their method is required in the literature. However, the focus of this article is not to examine the procedure of Molis and Chotiros[13] for measuring an effective grain bulk modulus, but to examine the use of such values for the unjacketed bulk modulus.

The range of compatible values between the unjacketed bulk modulus and frame bulk modulus is given by Eq. (5) above. Since Chotiros' analysis requires the assumption of an equivalent homogeneous solid matrix, then the measured value for the effective grain bulk modulus may be used for the unjacketed bulk modulus. Using Chotiros' ARL:UT values for the effective grain bulk modulus and porosity, the compatible range for the frame bulk modulus for this material is

$$0 < K_b^{ARL:UT} < 4.2 \times 10^9 \ \text{Pa}. \tag{12}$$

Chotiros' value of frame bulk modulus ($K_b = 5.3 \times 10^9$ Pa) is larger than the upper bound calculated using the value of effective grain bulk modulus ($K_r = 7.0 \times 10^9$ Pa). This suggests that under the current assumption the values of unjacketed bulk modulus and frame bulk modulus given by Chotiros[4] are incompatible. Therefore, one or both values are questionable or the assumption of an equivalent homogeneous solid is too restrictive. Further examination of the unjacketed bulk modulus is the most obvious choice since it is the one that has been kept constant in most previous models and is dependent on the assumption of an equivalent homogeneous solid.

Using the ARL:UT parameter set as input, the Biot theory is now used to predict the wave speeds for the two dilatational waves. Figure 1 shows the wave speed for two different values of unjacketed bulk modulus. The smaller value ($K_r = 7.0 \times 10^9$ Pa) is that of Chotiros,[4] where the incompatibility with the frame bulk modulus is ignored for the moment. The larger value ($K_r = 3.6 \times 10^{10}$ Pa) is that of quartz, which lies within the Voigt–Reuss–Hill bounds and satisfies the compatibility condition between the frame and unjacketed moduli. The effect of increasing the unjacketed bulk modulus is to increase the phase speed of the type I P wave while decreasing the phase speed of the type II P wave. This shows that small values of unjacketed bulk modulus are required to predict the larger phase speeds of the type II P wave that are observed by Chotiros.[4]

The phase velocity of the type I P wave predicted in Fig. 1 using the unjacketed bulk modulus of quartz is much larger than is measured by Chotiros.[4] Therefore, the frame bulk modulus has to be adjusted in order to fit the phase speed of the type I P wave. Figure 2 shows the phase speed
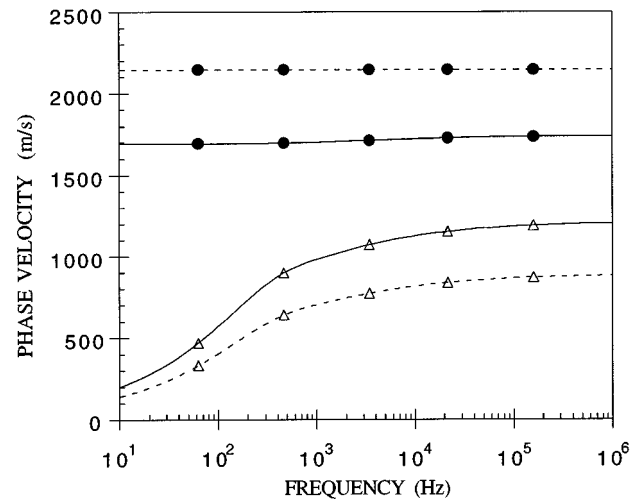
FIG. 1. The phase velocity of the type I (fast) P wave ● and type II (slow) P wave △ as predicted using the Biot theory for two values of "unjacketed bulk modulus." The solid lines are calculated using Chotiros' value, whereas the dashed lines are calculated using the value of quartz.

for two different values of frame bulk modulus. The values are that of Chotiros[4] ($K_b = 5.3 \times 10^9$ Pa) and the other ($K_b = 3.8 \times 10^8$ Pa) is chosen so as to fit the phase speed of the type I P wave. Reducing the frame bulk modulus reduces the phase speed of both the type I and type II P waves. By choosing a frame bulk modulus to fit the type I P wave speed in Fig. 2, we have essentially returned to predictions of previous models. These models predict a type II P wave having a small phase speed and very high attenuation. It can also be inferred that Chotiros'[4] larger value of frame bulk modulus is associated with his choice of a smaller value for the unjacketed bulk modulus.

From the above calculations, shown in Fig. 1 and Fig. 2, it is evident that in order to obtain the largest phase speed for the type II P wave, a minimum value of unjacketed bulk modulus, for a particular value of frame bulk modulus, is
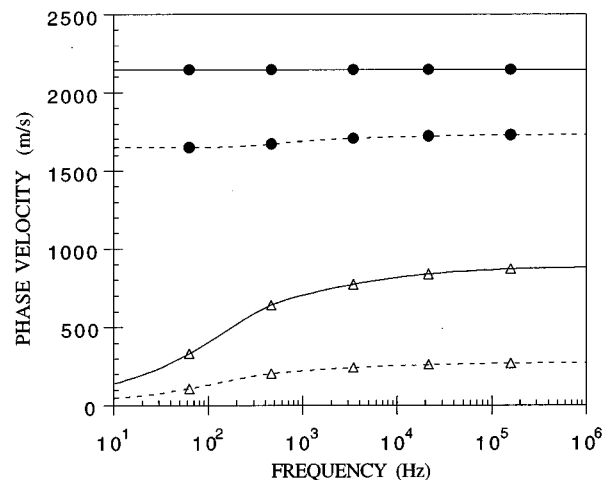
FIG. 2. The phase velocity of the type I (fast) P wave ● and type II (slow) P wave △ as predicted using the Biot theory for two values of "frame bulk modulus." The solid lines are calculated using Chotiros' value, whereas the dashed lines are calculated using the value of 3.8×10^9 Pa to fit the type I P wave speed of 1700 m/s. The unjacketed bulk modulus is that of quartz.
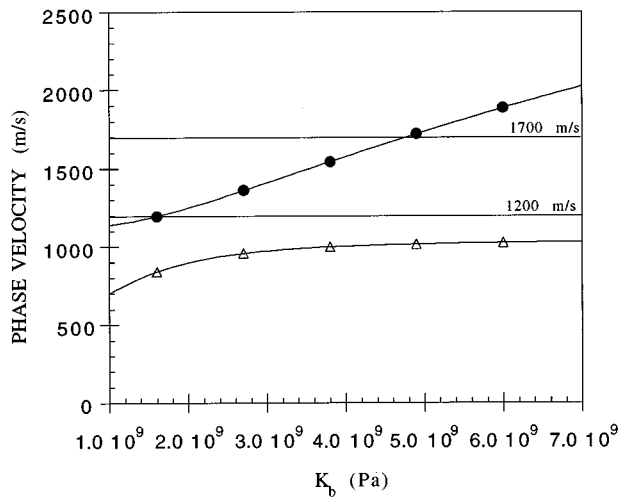
FIG. 3. The phase velocity of the type I (fast) P wave $\bullet$ and type II (slow) P wave $\triangle$ as a function of "frame bulk modulus." The "unjacketed bulk modulus" is calculated using the upper bound given in Eq. (5). Lines corresponding to approximate measured values are also placed on the graph.

FIG. 4. The phase velocity of the type I (fast) P wave $\bullet$ and type II (slow) P wave $\triangle$ as a function of "frame bulk modulus." The "unjacketed bulk modulus" is calculated using the upper bound given in Eq. (5). The "coefficient of fluid content" is calculated using Eq. (13) using $K_s = 7.0 \times 10^9$ Pa as measured by Molis and Chotiros.

required. In view of the measurements of Molis and Chotiros,[13] let's relax the stronger constraint on the unjacketed bulk modulus imposed by the Voigt–Reuss–Hill bounds, but still use Eq. (6). The constraint relating the frame bulk modulus, Eq. (5), to unjacketed bulk modulus, must be upheld. Figure 3 shows the calculated phase speed as a function of the frame bulk modulus for a frequency of 10 kHz. The unjacketed bulk modulus is calculated using the upper bound given by Eq. (5). Therefore, all three values, i.e., frame bulk moduli, unjacketed bulk moduli, and coefficient of fluid content are being varied. Chotiros' measured velocities are represented by the labeled horizontal solid lines in Fig. 3. The calculated phase speed of the type I P wave agrees with observed value of 1700 m/s for a frame bulk modulus of 4.8×10^9 and an unjacketed bulk modulus of 8.0×10^9 Pa. The predicted wave speed of the type II P wave is only 1013 m/s. This is about 20% less than the observed value. Furthermore, the phase speed of the type II P wave appears to have reached an asymptotic maximum. Based on the above, it is not possible to predict the observed phase speed of the type II P wave by only varying the unjacketed and frame moduli in a compatible fashion and calculating the coefficient of fluid content via Eq. (6).

It should also be noted that the value of unjacketed moduli used to construct Fig. 3 is a minimum value for compatibility. It's calculated from the upper bound of Eq. (5), which stems from the fact that the samples have infinite shear strength. This is quite contradictory to what one would expect because of the low shear wave speed. Therefore, from an intuitive viewpoint it appears that the unjacketed bulk modulus should be much larger than is used in Fig. 3, which will in turn make for a poorer prediction.

Relaxing the assumption of an effective homogeneous solid, one can no longer use Eq. (6). Hickey *et al.*[14] and Chotiros *et al.*[15] have shown that when the assumption of an equivalent homogeneous solid is relaxed the coefficient of fluid content can be calculated using
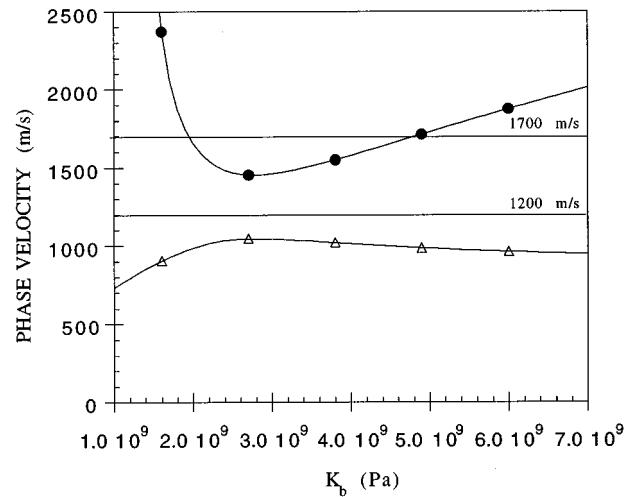
$$\frac{1}{K_r^f} = \beta \left(\frac{1}{K_f} - \frac{1}{K_r} \right) + (1 - \beta) \left(\frac{1}{K_s} - \frac{1}{K_r} \right). \tag{13}$$

The second term on the right-hand side of Eq. (13) represents the deformation associated with a change in porosity during an unjacketed test. It vanishes when the unjacketed bulk modulus is equal to the bulk modulus of the solid grains. To fully specify the problem, an additional macroscopic parameter is required. In the work of Chotiros,[4] the obvious choice is the effective bulk modulus of the solid grains. However, his parameter set now requires a measured value for the unjacketed bulk modulus *or* the coefficient of fluid content.

Figure 4 shows the wave speed as a function of frame bulk modulus. The unjacketed bulk modulus is calculated using the upper bound given by Eq. (5). Once the unjacketed bulk modulus is specified, the coefficient of fluid content is uniquely specified by using Eq. (13) with the solid bulk modulus measured by Molis and Chotiros.[13] The frequency is 10 kHz. For values of frame bulk moduli less than about 2.0×10^9 Pa, the phase velocity of the type I P wave begins an increase to unphysical values. The attenuation also behaves in an unphysical fashion. This suggests that there exist additional constraints on the values of frame, unjacketed, and grain bulk moduli. A constraint between unjacketed and effective grain bulk moduli is most likely required. This author is unaware of any such constraint in the literature.

The phase velocity of the type I P wave is equal to the observed value of 1700 m/s for two values of frame bulk modulus. These values are 1.94×10^9 and 4.85×10^9 Pa, with corresponding values of unjacketed bulk moduli of 3.23×10^9 and 8.10×10^9 Pa and values for the coefficient of fluid content of -2.18×10^{10} and 7.15×10^9 Pa. A negative value for the coefficient of fluid content is suspect but it is not unphysical. It simply means that during an unjacketed test the sample porosity is decreased sufficiently such that fluid exits rather than enters the sample. The phase velocity
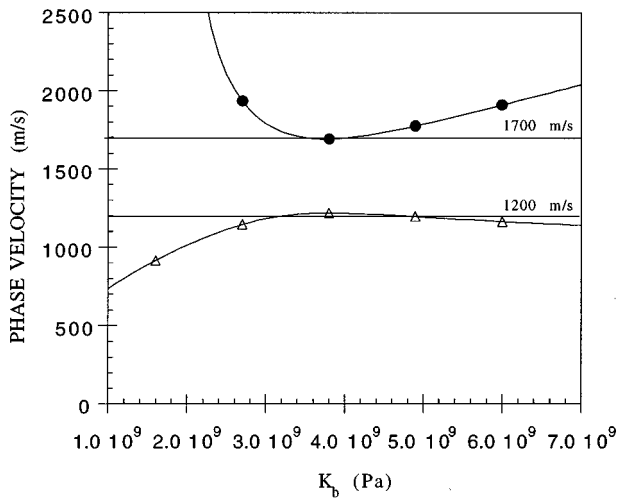
FIG. 5. The phase velocity of the type I (fast) P wave ● and type II (slow) P wave △ as predicted using Biot theory. Parameters are determined as in Fig. 4, except that the grain bulk modulus is that of quartz ($K_s = 3.6 \times 10^{10}$ Pa).

of the type II P wave never reaches the observed value of 1200 m/s.

Figure 5 shows the wave speed as a function of frame bulk modulus. The parameters are determined in the same fashion as for Fig. 4 except that the solid bulk modulus for quartz is used. The unphysical behavior of the type I P wave has extended to larger values of frame bulk modulus. Using this parameter set there is a range of values which could predict the phase velocities of both type I P wave and type II P wave as observed by Chotiros.[4] A type I P wave velocity of 1707 and type II P wave velocity of 1208 m/s is predicted for values of $K_b = 3.4 \times 10^9$, $K_r = 5.7 \times 10^9$, and $K_r^f = 5.56 \times 10^{10}$ Pa. A type I P wave velocity of 1707 and type II P wave velocity of 1212 m/s is also predicted for values of $K_b = 4.2 \times 10^9$, $K_r = 7.0 \times 10^9$, and $K_r^f = 1.94 \times 10^{10}$ Pa. It is quite interesting, but coincidental, that the value of unjacketed bulk modulus required to fit the data is the value of the effective grain modulus measured by Molis and Chotiros.[13] The attenuation predictions, however, do not agree with Chotiros' observations.

The above example calculations are far from a complete sensitivity analysis of the effects of relaxing the assumption of an equivalent homogeneous solid. It does show that the phase velocities observed by Chotiros[4] can be modeled using the Biot theory. However, the grain bulk modulus needed in this analysis is that of quartz, and not the value measured by Molis and Chotiros.[13]

III. CONCLUSION

A major problem with using Biot theory to model sound in ocean bottom sediments is obtaining values for several elastic parameters. The assumption of an equivalent homogeneous solid reduces the number of unknown independent parameters by requiring that the unjacketed bulk modulus equal the grain bulk modulus and that coefficient of fluid content be calculated by Eq. (6). A comparison of Biot theory with a theory developed via the volume averaging approach[16,17] has shown that the assumption of an equivalent homogeneous solid is a fundamental requirement for the Biot theory. However, such comparisons are currently under debate.

Chotiros[4] has detected transmitted (refracted) energy for angles below the type I P wave critical angle. He associates this energy with the type II (slow) P wave. In modeling his measurements, however, Chotiros[4] has used values of frame and unjacketed bulk moduli which are incompatible under the assumption of an equivalent homogeneous solid. It was shown here that it is not possible to predict Chotiros' observed large phase speed of the type II P wave by assuming an equivalent homogeneous solid. Relaxing the assumption of an equivalent homogeneous solid and using the grain bulk modulus measured by Molis and Chotiros[13] also does not appear to fit the data. However, it is possible to choose a compatible set of parameters for the Biot theory so as to fit the velocity observations. A set of independent measurements, including the unjacketed bulk modulus, grain bulk modulus, and coefficient of fluid content, are needed.

[1] M. A. Biot, "Theory of propagation of elastic waves in a fluid saturated porous solid: Low-frequency range," J. Acoust. Soc. Am. **28**, 168–178 (1956).
[2] M. A. Biot, "Theory of propagation of elastic waves in a fluid saturated porous solid: Higher frequency range," J. Acoust. Soc. Am. **28**, 179–191 (1956).
[3] M. A. Biot and D. G. Willis, "The elastic coefficients of the theory of consolidation," J. Appl. Mech. **24**, 594–601 (1957).
[4] N. P. Chotiros, "Biot model of sound propagation in water saturated sand," J. Acoust. Soc. Am. **97**, 199–214 (1995).
[5] A. Nur and J. D. Byerlee, "An exact effective stress law for elastic deformation of rocks with fluids," J. Geophys. Res. **76**, 6414–6419 (1971).
[6] L. W. Morland, "A simple constitutive theory for a fluid saturated porous solid," J. Geophys. Res. **77**, 890–900 (1972).
[7] J. R. Rice and M. P. Cleary, "Some basic stress diffusion solutions for fluid saturated elastic porous media with compressible constituents," Rev. Geophys. Space Phys. **14**, 227–241 (1976).
[8] R. J. S. Brown and J. Korringa, "On the dependence of elastic properties of a porous rock on the compressibility of the pore fluid," Geophysics **40**, 608–616 (1975).
[9] D. H. Green and H. F. Wang, "Fluid pressure response to undrained compression in saturated sedimentary rocks," Geophysics **51**, 948–956 (1986).
[10] J. G. Berryman and G. W. Milton, "Exact results for generalized Gassmann's equations in composite porous media with two constituents," Geophysics **56**, 1950–1960 (1991).
[11] J. G. Berryman, "Effective stress for transport properties of inhomogeneous porous rock," J. Geophys. Res. **97**, 17,409–17,424 (1992).
[12] R. Hill, "Elastic properties of reinforced solids: Some theoretical aspects," J. Mech. Phys. Solids **11**, 357–372 (1962).
[13] J. C. Molis and N. P. Chotiros, "A measurement of grain bulk modulus of sands," J. Acoust. Soc. Am. **91**, 2463 (1992).
[14] C. J. Hickey, T. J. T. Spanos, and V. de la Cruz, "Deformation parameters of permeable media," Geophys. J. Int. **121**, 359–370 (1995).
[15] N. P. Chotiros, J. C. Molis, and J. L. Darrouzet, "A measurement of the modulus of sand grains," J. Acoust. Soc. Am. (submitted).
[16] V. de la Cruz and T. J. T. Spanos, "Seismic wave propagation in a porous medium," Geophysics **50**, 1556–1565 (1985).
[17] V. de la Cruz, P. N. Sahay, and T. J. T. Spanos, "Thermodynamics of Porous Media," Proc. R. Soc. London, Ser. A **243**, 247–255 (1993).

Attenuation in suspensions of irregularly shaped sediment particles: A two-parameter equivalent spherical scatterer model

Arjen S. Schaafsma
Delft Hydraulics, P.O. Box 177, 2600 MH Delft, The Netherlands

Alex E. Hay
Department of Oceanography, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada

Results are presented from an acoustic attenuation spectroscopy study of suspensions of irregularly shaped sediment particles and of nominally spherical lead-glass beads. The measured spectra cover a wide frequency band of 1–100 MHz. It is shown that the spectra can be brought into close agreement with the rigid movable sphere model using two equivalent sphere size parameters. One parameter is the diameter d_G of an equal cross-sectional area sphere, and is used to scale the frequency to obtain the nondimensional size/frequency parameter ka. The other parameter is the diameter d_p of an equal volume sphere, and is used to scale the particle volume to obtain the particle number density N. It was observed that for the irregular particles the attenuation for $ka > 1$ is enhanced with respect to the sphere case. The observed enhancement factors are considerable: up to 1.6 for natural sand particles and 1.8 for the most irregular particles studied (ground quartz). In terms of the two-parameter model the enhancement factor is equal to a shape parameter $b_0 = (d_G/d_p)^3$ and can indeed be explained as a shape effect. The implications of this result for the acoustic detection of suspended sediments in aqueous environments are discussed. © 1997 Acoustical Society of America. [S0001-4966(97)00909-0]

LIST OF SYMBOLS

a	particle radius
a_G	radius of equal cross-sectional area sphere
a_m	measured particle radius
a_p	radius of equal volume sphere
a_0	scaling parameter for measured particle radius
b_0	scaling parameter for particle volume
C_m	mass concentration of particles
C_v	volume concentration of particles
d	particle diameter
d_s	average sieve diameter
d_y	diameter at $y\%$ of cumulative particle size distribution by volume
G_p	geometric cross-sectional area of particle
k	acoustic wave number
N	number density of particles
V_p	particle volume
V_G	volume of equal cross-sectional area sphere
x	ka, nondimensional size/frequency parameter

Greek

α	linear attenuation constant for pressure
λ	wavelength
ρ	density of particle
σ	standard deviation
Σ	total attenuation cross section

INTRODUCTION

There has been considerable interest in the development of acoustic methods for sediment transport measurement during recent years.[1–16] The main attraction of the acoustic approach is that it has the potential of providing vertical profiles of suspended sediment concentration and size remotely, and therefore with negligible flow disturbance at the bed, with high spatial and temporal resolution. The most direct way to obtain these profiles would be from backscattering measurements, as in acoustic current profiling.[17] This is the reason that most of the interest referred to above has been in backscattering systems. However, the interpretation of the backscattered signals is a rather complex inversion problem, which is far from being solved completely.

The measurement of the frequency dependence of the attenuation (or extinction), has also received some attention.[6,18–21] This method cannot provide profiles, but it has been shown that it can provide accurate local measurements of the concentration and size in the case of spherical particles.[21] The present paper is a further step in the development of this method, by extending the analysis to irregularly shaped sediment particles, following a semi-empirical approach.

Very few measurements have been made of the scattering cross sections of natural sediment particles. In particular, there have been few measurements made of the total scattering cross section of suspended sand. With respect to the backscattering inversion problem, the total scattering cross section appears in an exponential term (the attenuation due to scattering) in the equation for the backscattered signal amplitude. Therefore the estimated concentrations are rather sensitive to errors in this term.

In this paper results are presented from a series of experiments in which the total scattering cross section was measured in a frequency-swept acoustic attenuation spectrometer, operating from 1 to 100 MHz, for natural sand particles and for lead-glass beads. It is noted that some preliminary results have been presented,[22,23] since the experimental data became available.[24]

The relation between the total cross section Σ and the attenuation constant α is, by definition,

$$\Sigma = \frac{2\alpha}{N}, \qquad (1)$$

where N is the number concentration of particles in suspension. In principle it is possible to determine Σ experimentally by measuring α for known values of N. The measured cross sections can then be compared with theoretical calculations, using a suitable estimate of the scatterer radius a to normalize the acoustic wave number k. For particles in suspension, it is usually not possible to measure N directly. Instead the volume concentration C_ν of the particles is determined, often from a gravimetric measurement of the mass concentration in a sample of the suspension and the known bulk density of the particles. Then N is given by

$$N = \frac{C_\nu}{V_p}, \qquad (2)$$

where V_p is the particle volume. If the particle volume cannot be measured directly, it is usually estimated from a measurement of the particle diameter. For spherical particles, the exact relation

$$V_p = \frac{4\pi}{3} a^3 \qquad (3)$$

can be used. It is clear that a small error in the estimate of a will lead to a comparatively large error in V_p, and therefore in N. For irregular particles the errors in these estimates are likely to be larger, since the above equation can be used in an effective sense only, and supposes that the relevant (equivalent volume) effective size can be measured.

Consider now the problem of sound scattering in a turbulent suspension of irregularly shaped particles. Because the particles have random shapes and adopt random orientations as a function of time due to the turbulent motion, one might expect that the time-averaged behavior of the scattered sound field could be described, to a reasonable first approximation, by assuming a spherical shape for the scatterers and some effective size. The few available measurements of the total scattering and backscattering cross sections of natural sand particles support this assumption.[6,16,18,20,25–27] The choice of effective size for the equivalent sphere for natural sand grains has not yet been examined, however. It is this question which is investigated here.

It is clear from Eqs. (1)–(3) that two effective sizes for the equivalent sphere are involved: one is required to scale the wavelength to obtain ka, the other to scale the particle volume to obtain N. To illustrate, the left-hand side of Eq. (1) can be computed theoretically, for any value of the non-dimensional size/frequency parameter ka. Comparison with the experimentally determined quantity on the right-hand side then requires a suitable choice of effective radius with which to normalize the values of k at which the measurements of α were made. The natural choice for this effective

size corresponds, for measurements of the total scattering cross section, to the average geometric cross-sectional area of the particles. The second effective particle size corresponds to an equivalent sphere with the same volume V_p as the particle, and is required to convert the measured volume concentration of particles to the number concentration N. One effective size therefore corresponds to the diameter of a circle of area equal to the projected area of the particle, the other to the diameter of an equal volume sphere. These two sizes will not in general be the same. It may therefore be anticipated that there should be two different renormalization scales involved in making comparisons between total scattering cross-section measurements for irregular particles, and theoretical calculations for spherical scatterers.

I. EXPERIMENTAL METHODS

A. Laboratory measuring system

The experimental setup used to perform the acoustic attenuation measurements in the frequency range of 1–100 MHz is shown schematically in Fig. 1. The main features can be summarized as follows. The suspension to be measured is contained in a circuit and circulated by means of a slurry pump. Measuring sections are mounted in a vertical perspex cylinder (inner diameter 10 cm). A diffuser at the top of the vertical cylinder homogenizes the particle suspension. Samples of the suspension can be taken by a suction tube, used to obtain reference values for the mass concentration and the particle size distribution. The temperature of the water is regulated to within ± 0.01 °C. This is necessary because the temperature dependence of the attenuation is large at higher frequencies, about 80 dB/m per °C at 100 MHz.[28]

The attenuation measurements are carried out with six pairs of diametrically opposed broadband acoustic transducers, placed in two measuring sections. Basically, measurements are taken at a single frequency and the frequency is swept to obtain a spectrum. The signal generation and processing are explained briefly, referring to the block diagram in Fig. 1. Of each transducer pair, one transmits and the opposite one receives. The transmitted signal is a single-frequency sine wave burst. The burst duration is relatively short with respect to the acoustic path length in water, in order to avoid reflections and interference effects. Typical burst lengths are 15 μs below 20 MHz and 2 μs above. The average amplitude of the received burst signal is determined, digitized, and stored. An average over a large number of bursts (typically 1000) is taken to increase the measurement accuracy. The burst repetition rate is 1 kHz. At each frequency an electronic reference signal is measured, stored, and subtracted from the detected acoustic signal. This partly eliminates the drift and the frequency dependence of the electronic system. The center frequencies of the transducer pairs are 2.25, 4.5, 10, 30, 50, and 100 MHz. The transducers are commercially available, broadband, immersion-type transducers (from Panametrics, except the 4.5-MHz transducers). Their bandwidths overlap partially, so that depending upon frequency, measurements can be taken with one or more pairs. The separations between the transducer pairs are
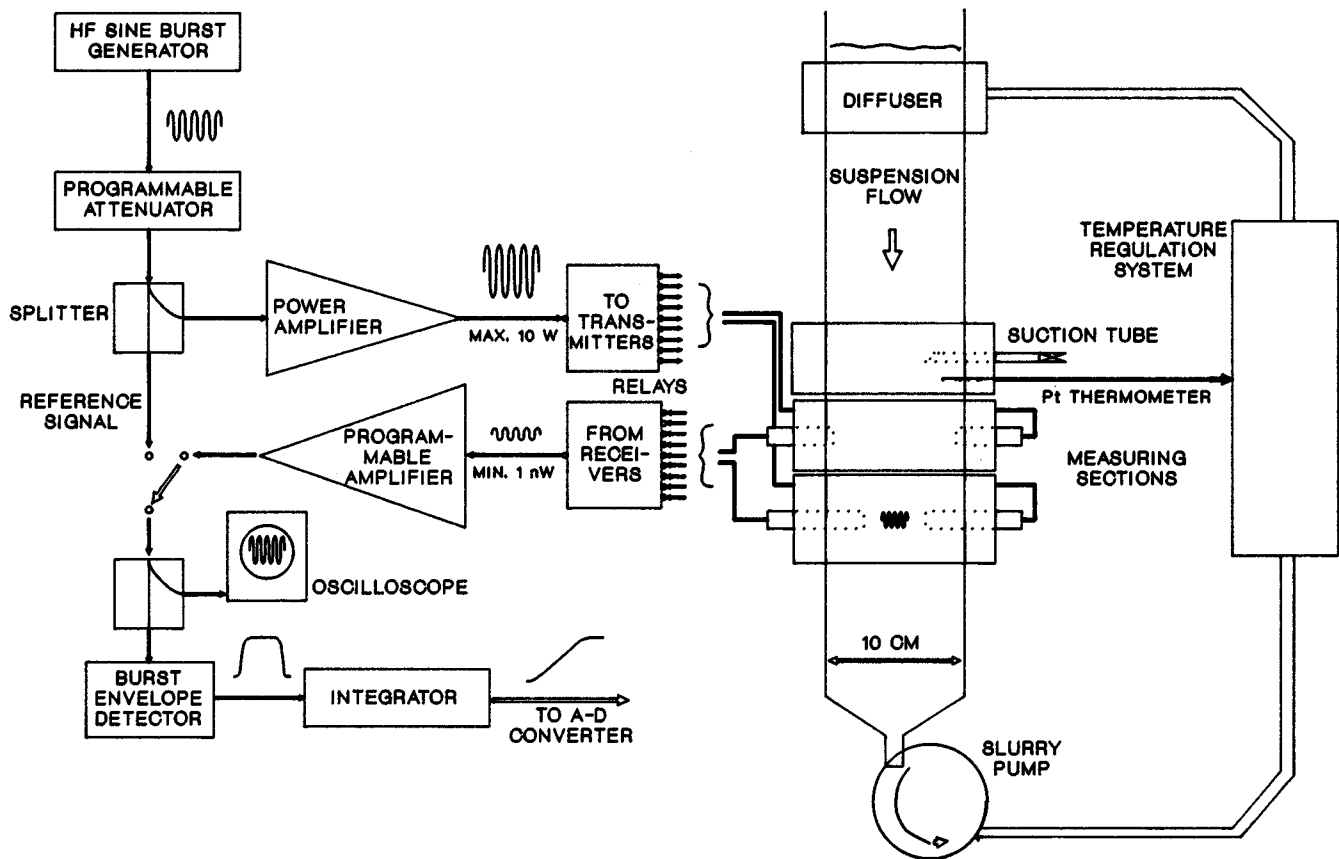
FIG. 1. Sketch of the experimental setup, showing a block diagram of the electronics and the suspension circuit.

given in Table I along with other characteristic properties.

The attenuation spectrum of suspended particles is measured relative to pure water. That is, attenuation data are first collected for pure water, in the above described way, for each transducer pair, at a chosen number of frequencies. These data represent the sum of the electronic and acoustic attenuation, with respect to some fixed reference signal level. Second, a similar set of data is collected for the particle suspension. The difference between the second and the first set of data is the acoustic attenuation spectrum of the suspended particles. Typically, it takes about 1 h to collect a complete data set.

TABLE I. Characteristic parameters of the six transducer pairs used for the attenuation measurements: the center frequency f_0, the effective transmission bandwidth BW, the diameter of the piezoelectric element D_t, the near-field distance in water r_0 (at f_0), and the transducer separation l. The near-field distance is defined as the region where the beam is approximately cylindrical [see, e.g., L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics* (Wiley, New York, 1962), 2nd ed., Chap. 7.12]: $r_0 = 1.64 D_t^2/(4\lambda_0)$.

f_0 (MHz)	BW (MHz)	D_t (mm)	r_0 (mm)	l (mm)
2.25	1–3.5	12.7	100	100
4.5	2.4–5.3	5.0	31	100
10	4.9–15	6.35	112	72
30	11.5–34	6.35	336	100
50	12–50	3.18	140	50
100	24–100	3.18	280	28

B. Validation experiments

The attenuation as defined by Eq. (1) can be measured directly by determining the intensity decay of the acoustic wave in its propagation direction over a known distance. This is straightforward for well-defined wave patterns, like a plane wave or a spherical wave. However, for the present measuring system the transducer configurations are such that most or all of the acoustic path length is within the near-field zone of the transducers (see Table I). It is well-known that in this zone the beam pattern of a transducer has a rather complex structure and is approximately nondiverging. It is expected that, because the transmit and receive transducers used have the same area, an averaging takes place over the structure in the near-field beam pattern. Therefore it is expected that changes in the measured attenuation should represent the average attenuation in the space separating the transducers. Nevertheless, it is worthwhile to verify the influence of near-field effects experimentally. For this purpose, two types of experiments were carried out.

In the first place, the change of the attenuation spectrum of pure water with temperature was measured and compared with results available in the literature. The results of these relative measurements are shown in Fig. 2. The change of the attenuation constant α, relative to its value at 19.8 °C, was measured for four values of the temperature difference: -1.0, -0.5, $+0.5$, and $+1.0$ °C. Also shown in the figure are continuous curves, which correspond to the empirical equation given by Fisher and Simmons [Ref. 28, Eq. (10)].
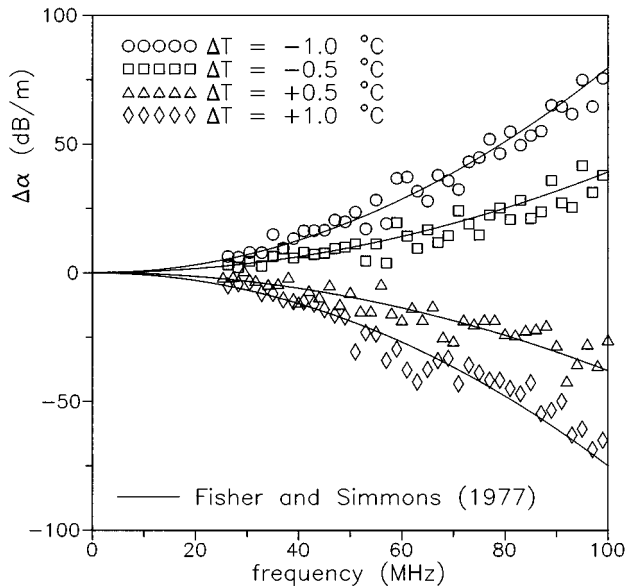
FIG. 2. Measured temperature dependence of the attenuation, relative to its value at 19.8 °C, compared with the empirical equation given by Fisher and Simmons, corresponding to the solid curves (Ref. 28). To improve readability, the data between 1 and 25 MHz, which are close together on the solid curves, are not shown and above 25 MHz only half of the data points are plotted.

This equation agrees to within $\pm 4\%$ with the experimental data of Pinkerton,[29] which are in the MHz range up to about 70 MHz. Pinkerton verified that his measurements were not affected by near-field effects, by carrying out separate measurements in the near-field as well as in the far-field zones of the transducers. It is shown in Fig. 2 that the present experimental data agree very well with those of Pinkerton. The agreement is within $\pm 10\%$ (standard deviation) over the whole frequency range.

The second type of experiments involved a detailed comparison between measured and calculated attenuation spectra for spherical lead-glass particles. The result was a good quantitative agreement, the discrepancy was maximum 10% in the high end of the frequency range and smaller in the remaining part.[21]

The above results give confidence that an attenuation

spectrum as measured with the present system, is accurate to within approximately 10% and that effects related to the near-field geometry are not important within this error.

C. Characterization of suspensions and particles

For a discussion of the acoustic measurements in terms of suspension concentration, particle size, and shape, an independent reference measurement of these parameters is needed. The measurement of the suspension concentration using the suction system of the experimental setup has been developed into a reliable and accurate method. As discussed in some detail before, an accuracy of 5% (standard deviation) has been realized for the mass concentration measurement.[21] In the laboratory experiments suspension concentrations representative for field conditions were used, up to a mass concentration of 5 kg/m^3, corresponding to a maximum concentration by volume of 0.19%. Because of these low concentrations it is unlikely that multiple-scattering effects are important.

Three sets of particle size fractions were prepared for the present study by sieving or by the use of a hydrocyclone. A reference for the particle size has been obtained by a standard optical diffraction analysis method, using a commercial instrument (Malvern particle sizer 3600 D). Some characteristic results of the particle size analysis for the first set of size fractions, consisting of natural sediment and ground quartz particles, are given in Table II. Average sizes range from about 200 μm down to 30 μm. A second set of size fractions, in approximately the same size range, was made out of nominally spherical lead-glass beads and size analysis results are given in Table III. These particles were chosen for the purpose of comparison with the irregularly shaped sediment particles. Finally, a third set of sediment size fractions was prepared out of three natural sediments of a different origin in two subsets of average size 100 and 200 μm, respectively (see Table IV). In this case also the so-called sphericity was determined, a geological classification by optical microscopy methods.[30]

The densities of the particles are needed to convert measured concentrations by mass to concentrations by volume: $C_v = C_m/\rho$. Densities were measured by a volumetric

TABLE II. Summary of optical diffraction (OD) size analysis data and of the two-parameter model interpretation of the acoustic attenuation spectra for the sand sieve size fractions and the quartz hydrocyclone fraction (HC). Assuming a Gaussian size distribution, the standard deviation σ_m was calculated from the OD sizes. The semiempirical analysis was carried out for a single size ($\sigma_{th}=0$) and in four cases also for a Gaussian size distribution, with σ_{th} close to the experimental value. Apart from the indicated exceptions, all measurements were carried out at 19.8 °C.

Particles	Size fraction (μm)	OD sizes				2-p model			2-p model		
		d_{16}	d_{50} (μm)	d_{84}	σ_m (%)	σ_{th} (%)	a_0 (−)	b_0	σ_{th} (%)	a_0 (−)	b_0
Dune sand	180–212	176	212	288	26	0	0.80	1.14	30	0.80	1.18
Assen	90–106	86	110	137	24	0	0.71	1.46	20	0.72	1.48
Assen	45–53	43	60	79	30	0	0.71	1.56	30	0.75	1.61
Assen[a]						0	0.68	1.59			
Assen	45–53	44	62	82	31	0	0.67	1.59			
Assen[b]						0	0.62	1.53			
Quartz	HC4	21	30	45	40	0	0.63	1.82	40	0.60	2.05

[a]Measurement at 9.9 °C.
[b]Measurement at 29.7 °C.

TABLE III. Summary of optical diffraction (OD) size analysis data and of the two-parameter model interpretation of the acoustic attenuation spectra for the sieve size and hydrocyclone (HC) fractions of lead-glass beads. Assuming a (split-)Gaussian size distribution, the standard deviation(s) σ_m was (were) calculated from the OD sizes. The semiempirical analysis was carried out for a single size ($\sigma_{th}=0$) and in most cases also for a (split-)Gaussian size distribution, with a σ_{th} close to the experimental one. Apart from the indicated exceptions, all measurements were carried out at 19.8 °C.

Size fraction (μm)	OD sizes				2-p model			2-p model		
	d_{16}	d_{50} (μm)	d_{84}	σ_m (%)	σ_{th} (%)	a_0 (−)	b_0	σ_{th} (%)	a_0 (−)	b_0
180–212	>168	196	<225	<15	0	0.99	1.05	10	0.99	1.07
90–106	85	104	131	22	0	0.94	1.15	20	0.95	1.17
45–53	27	40	50	30	0	0.88	1.44	30	0.90	1.48
45–53[a]					0	0.88	1.28			
45–53	36	46	60	26	0	0.87	1.24			
45–53[b]					0	0.85	1.28			
HC4	18	23	43	54	0	0.89	1.40	22/87	0.78	1.59
HC5	12	17	30	53	0	0.95	1.49	29/76	0.71	1.89

[a]Measurement at 11.0 °C.
[b]Measurement at 29.6 °C.

method for the individual size fractions, as a check for homogeneity. For sand and quartz all values were within 1% of the average value of 2650 kg/m^3. For the lead-glass beads, a value of 2870 kg/m^3 was found for a bulk piece of the same lead-glass and for all size fractions, except for the two smallest sizes (obtained by a hydrocyclone). In the latter cases the densities were 2520 and 2495 kg/m^3, for the HC4 and HC5 size fractions, respectively (cf. Table III). Possible reasons for this inhomogeneity are discussed below.

A qualitative characterization of particle shape was carried out by scanning electron microscopy (SEM). Typical SEM micrographs for the three sets of size fractions are shown in Figs. 3–5, respectively. The sand particles from the three size fractions obtained by sieving, have an increasingly irregular shape the smaller their average size. This trend is continued for the quartz particles from the hydrocyclone size fraction. One can describe the general shape of these sediment particles as irregular polyhedra, with a varying number of faces and a varying degree of rounding (smoothing) of the edges. In this picture, the 200-μm sand particles are like polyhedra with a relatively large number of faces and relatively rounded edges. On the other, more irregular, end of this scale are the quartz particles (30-μm size fraction), which are closer to polyhedra with a relatively small number

of faces (e.g., a tetrahedron) and further relatively sharp edges.

The lead-glass beads are quite spherical on the whole, but most have surface and volume defects, some particles have a distorted shape. The surface defects are scratches and pits, the volume defects are interior holes, which can be concluded from an examination of the cross sections of the beads, as in a previous study.[21] The number of these defects appears to increase with decreasing average size of the beads, but cannot be quantified accurately. It is noted that a large number of internal cavities could explain the lower measured densities for the two smallest size fractions, mentioned above. These densities are lower with respect to the bulk value by 12% and 13% for the 23- and 17-μm size fractions, respectively. However, an alternative explanation for the lower densities is that these particles have a somewhat different composition: that is, a lower content of lead. It is not unlikely that this can happen in the melt spray production process of the lead-glass beads.

The SEM micrographs for the three 100-μm sand fractions of different origin (Fig. 5) show subtle differences in agreement with the geological sphericity classification (Table IV). In particular, the Assen sand particles appear

TABLE IV. Summary of size and acoustic analysis, as in Table II, for 100- and 200-μm size fractions of some different natural sands. Also given is the result of a geological classification by optical methods, the so-called sphericity, to be compared with the acoustic shape parameter b_0.

Particles	Size fraction (μm)	OD sizes				2-p model			Geological sphericity
		d_{16}	d_{50} (μm)	d_{84}	σ_m (%)	σ_{th} (%)	a_0 (−)	b_0	
Dune sand	180–212	176	212	288	26	0	0.80	1.14	middle
Twente	180–212	171	208	283	27	0	0.79	1.10	high
Ottawa	180–212	178	212	283	25	0	0.81	1.11	middle to high
Assen	90–106	86	110	137	24	0	0.71	1.46	low
Twente	90–106	93	115	142	21	0	0.73	1.32	high
Ottawa	90–106	91	115	143	23	0	0.74	1.28	middle to high
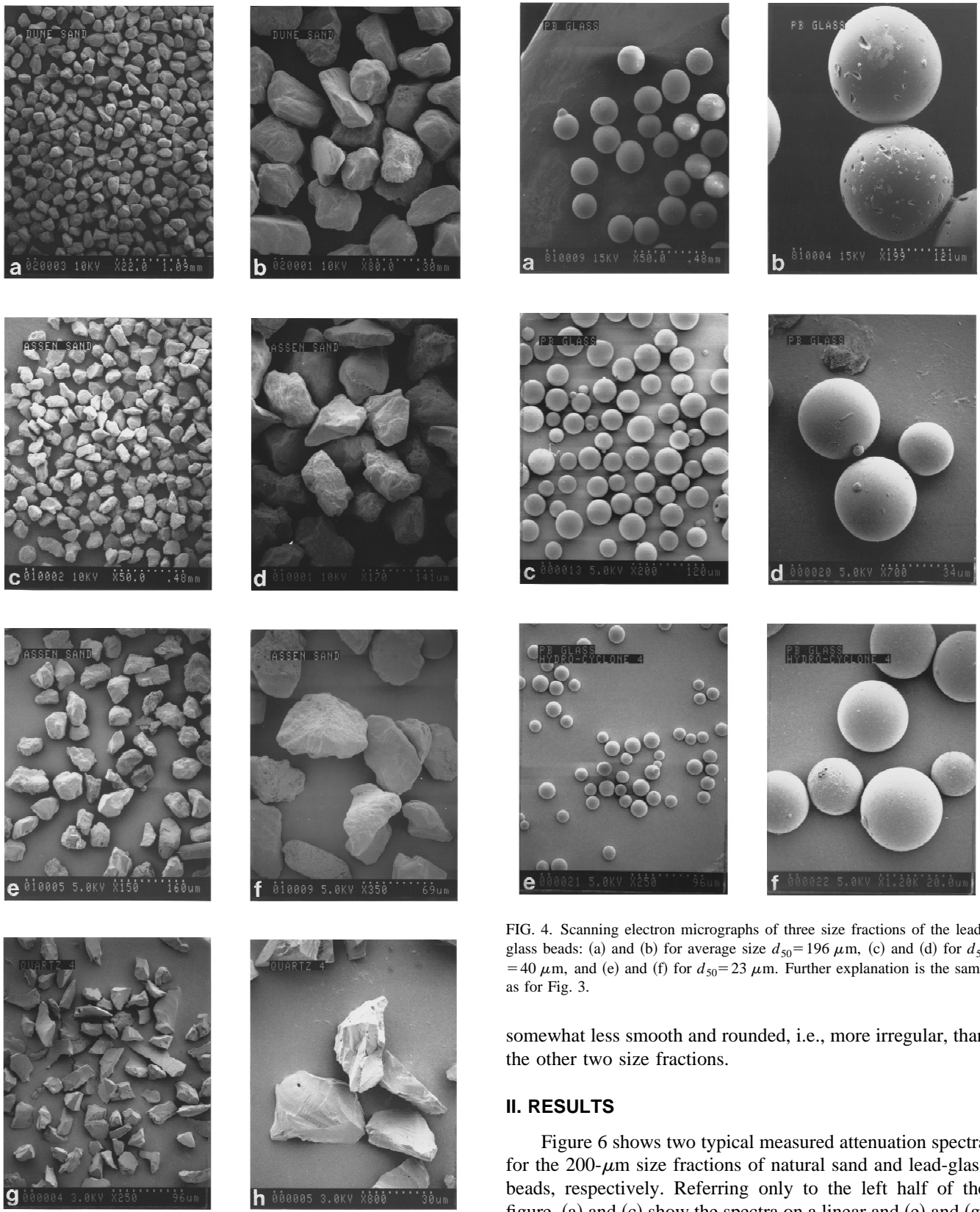
FIG. 3. Scanning electron micrographs of four size fractions of the sand and quartz particles, each at two different magnifications given in the lower right-hand corner of the micrographs together with a size scale. From top to bottom: (a) and (b) dune sand of average size $d_{50}=212$ μm, (c) and (d) Assen sand with $d_{50}=110$ μm, (e) and (f) Assen sand with $d_{50}=60$ μm, and (g) and (h) quartz particles with $d_{50}=30$ μm.

FIG. 4. Scanning electron micrographs of three size fractions of the lead-glass beads: (a) and (b) for average size $d_{50}=196$ μm, (c) and (d) for $d_{50}=40$ μm, and (e) and (f) for $d_{50}=23$ μm. Further explanation is the same as for Fig. 3.

somewhat less smooth and rounded, i.e., more irregular, than the other two size fractions.

II. RESULTS

Figure 6 shows two typical measured attenuation spectra for the 200-μm size fractions of natural sand and lead-glass beads, respectively. Referring only to the left half of the figure, (a) and (c) show the spectra on a linear and (e) and (g) on a logarithmic ka scale. The attenuation is presented as $\alpha/(kb_0C_v)$, where $b_0=1$, a normalization which will be explained below. The wave numbers corresponding to the measurements have been nondimensionalized using the median radii determined by optical diffraction ($d_m/2$), as given in Tables II and III, noting that $d_m=d_{50}$. Also shown in the figure is the theoretical attenuation for an acoustically rigid,
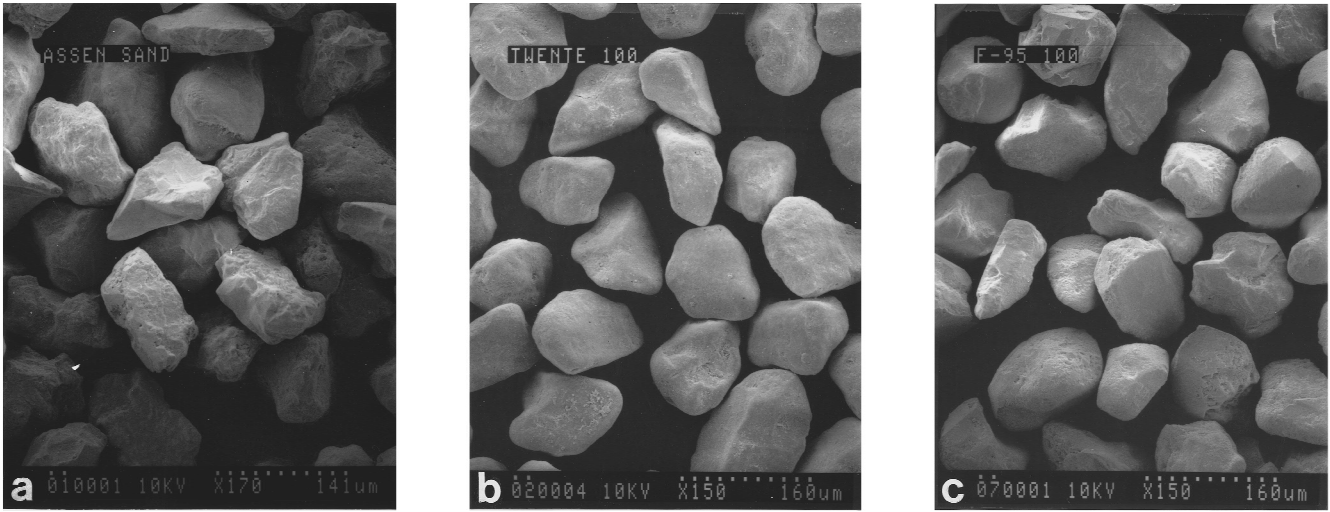
FIG. 5. Scanning electron micrographs of the 90–106-μm sieve size fractions of three sands of different sphericity: (a) Assen, (b) Twente, and (c) Ottawa sand (see Table IV). Further explanation is the same as for Fig. 3.

but mechanically movable sphere with the density of sand, in the same plots as the sand experimental data. Similarly, for lead-glass, the theoretical attenuation for an elastic lead-glass sphere is shown.[21] It can be seen that in the case of the lead-glass beads there is close agreement between the calculated and measured spectra. Also it is seen that for the sand particles the shape of the measured spectrum is very similar to that calculated theoretically, in particular resonance features are absent in the spectra of these nonspherical particles. However, the measured peak attenuation is somewhat greater than predicted by the rigid sphere model and occurs at a somewhat higher value of ka. It will now be demonstrated that this observed discrepancy in the case of the irregularly shaped sand particles, can be attributed to the scaling problems, related to the choice of effective size, discussed in the Introduction.

Let G_p be the average geometric (i.e., projected) cross-sectional area of the particles. The logical choice of scale for the total scattering cross section, in the high-frequency limit, is $2G_p$. Defining an equivalent circular radius a_G such that

$$\pi a_G^2 = G_p \tag{4}$$

then Eq. (1) becomes

$$\frac{\Sigma(x)}{2\pi a_G^2} = \frac{V_p}{G_p}\frac{\alpha(x)}{C_\nu} \tag{5}$$

in which Eq. (2) has been used to eliminate N, and $x = ka_G$. Though the further treatment could have been based on Eq. (5), it has several advantages to remove the explicit size dependence from the right-hand side of this equation, while preserving its nondimensional form, by dividing by $4ka_G/3$, yielding

$$\frac{3\Sigma(x)}{8ka_G\pi a_G^2} = \frac{V_p}{V_G}\frac{\alpha(x)}{kC_\nu} \tag{6}$$

in which now the ratio of the actual particle volume V_p to the volume of the equal cross-sectional area sphere $V_G = 4\pi a_G^3/3$ appears on the right-hand side. This ratio is equal to unity for a sphere and this is the main advantage to use

Eq. (6), instead of (5), for an analysis of deviations from the sphere case. Making the substitution $x = ka_G$, the above equation can be rewritten as

$$\frac{3\Sigma(ka_G)}{8ka_G\pi a_G^2} = \frac{\alpha(a_0ka_m)}{kb_0C_\nu} \tag{7}$$

where the nondimensional parameters a_0 and b_0 have been defined such that

$$a_0 = \frac{a_G}{a_m} \tag{8}$$

where $2a_m$ is the measured particle diameter (here, by optical diffraction) and

$$b_0 = \frac{V_G}{V_p}. \tag{9}$$

It can be seen from Eq. (7) that a_0 and b_0 are scaling parameters, a_0 scaling the measured sizes, and b_0 the measured concentrations. Again, the existence of the two parameters a_0 and b_0 reflects the two scaling problems involved in approximating irregularly shaped particles by spherical scatterers: one being the size required to nondimensionalize the wavelength, the other the size required to determine N from measurements of C_ν.

Because of the choice of a_G to scale the acoustic wavelength, so that $x = ka_G$, the left-hand side of Eq. (7) is readily computed from the theory for a spherical scatterer. Denoting the left-hand side by $F_{\text{th}}(x)$,

$$F_{\text{th}}(x) = \frac{3\Sigma(x)}{4x2\pi a^2} = \frac{3}{4x^2}\,\text{Im}[f_\infty(x,0)], \tag{10}$$

where $f_\infty(x,0)$ is the forward scattering form factor. Thus it is seen that the choice of the geometric cross section as the appropriate scale for the total scattering cross section leads in a straightforward way to a_G being the most suitable measure of size with which to nondimensionalize the acoustic wavelength. Note that dividing Eq. (5) by x leads to a $1/x^2$ dependence in Eq. (10). As will be seen, this scaling gives the

A. S. Schaafsma and A. E. Hay: Suspension attenuation   1491
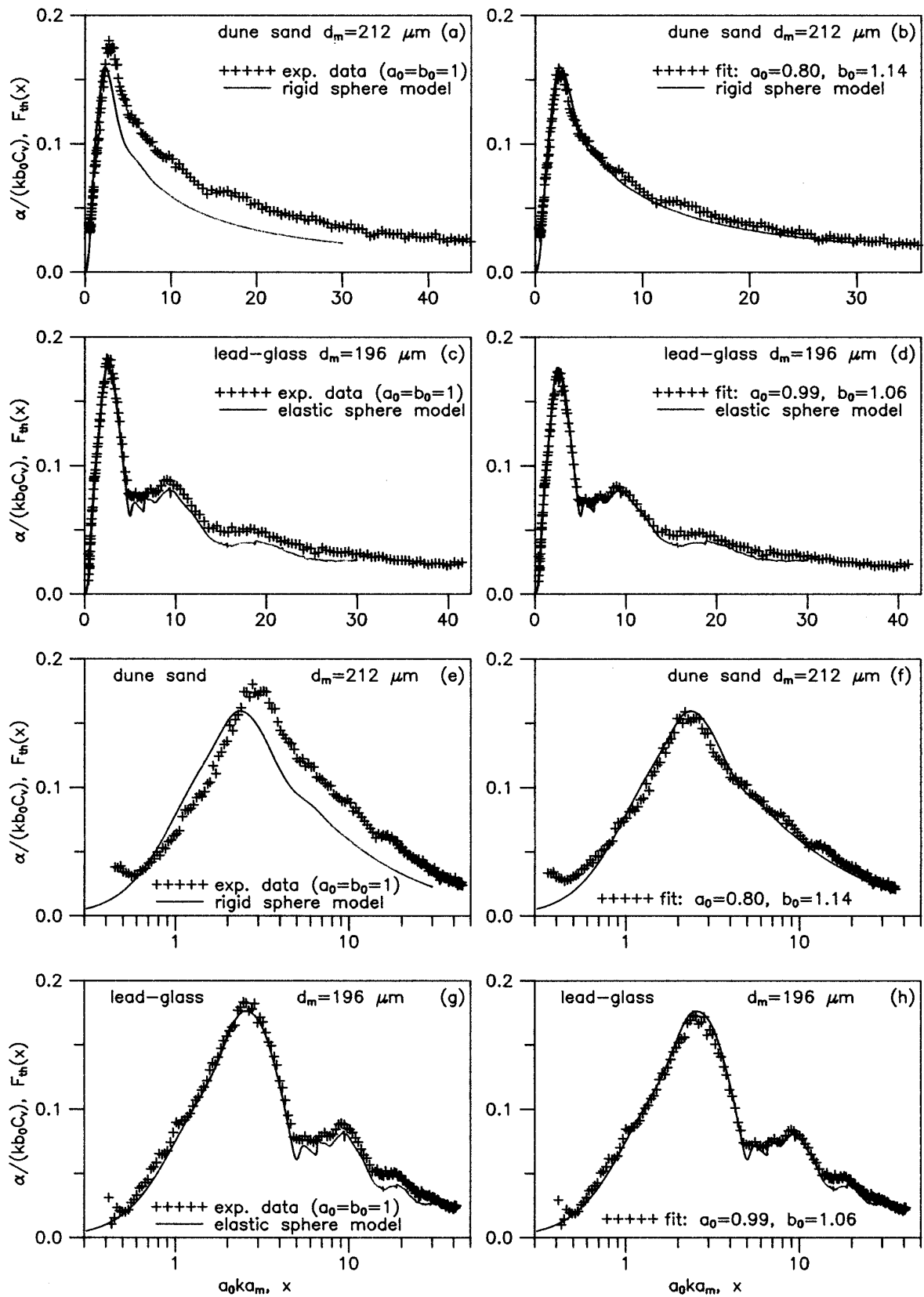
FIG. 6. Comparison of the measured normalized attenuation spectra of the 180–212-μm sieve size fractions of sand and lead-glass beads with the appropriate sphere models (rigid movable quartz and elastic lead-glass, respectively, for which the calculations were made until $ka=30$). In the left half of the figure the experimental data points are given as they were measured ($a_0=b_0=1$). In the right half of the figure, the measured data have been fitted to the model curves, using a two-parameter fitting procedure. The only difference between the plots (a)–(d) in the upper half of the figure and plots (e)–(h) in the lower half is the normalized frequency scale, which is linear and logarithmic, respectively. Note that there is a significant discrepancy between the as-measured data and the theoretical model for the sand particles, whereas for the lead-glass beads there is a good agreement.

greatest weight to the attenuation measurements in the neighborhood of $x=2$, that is where the main broad peak of the spectra is located, which is not unreasonable, since this is the region of the spectrum which provides the size sensitivity. This is seen easily if it is realized that when the normalized spectra are plotted versus frequency f, the peak position is size dependent (inversely proportional to the size).

The right-hand side of Eq. (7) can be determined from the measurements. That is,

$$F_{\text{exp}}(x) = \frac{\alpha(x=a_0 ka_m)}{kb_0 C_\nu} \tag{11}$$

for given values of a_0 and b_0. Here the values of these parameters are the unknowns, and are estimated for each experiment with particles of a given size by least-squares fitting the measurements to the theory for a rigid movable spherical scatterer. This is done by minimizing

$$\chi^2 = \frac{\Sigma_{j=1}^n (F_{\text{exp},j} - F_{\text{th},j})^2}{(n-2)\sigma^2}, \quad n>2, \tag{12}$$

where σ is the standard deviation of the measurements, n is the number of frequencies used in the fit, and x_j is the value of x at the jth frequency. The standard deviation of the measurements σ has been assumed equal for all data points, since the experimental error in the attenuation α increases approximately linear with frequency.

The results of applying this procedure to the data in the left half of Fig. 6 are presented in the right half of the same figure. The values of a_0 and b_0 are 0.80 and 1.14 for the 200-μm sand particles. It is important to note that choosing a_0 different from unity results in a rescaling of the values of x corresponding to the measurements only, and therefore in a displacement of the measured points parallel to the horizontal axis, with the theoretical curve fixed. Similarly, choosing b_0 different from unity amounts to rescaling C_ν only, and results in a vertical displacement of the measured points leaving the theoretical curve unchanged. For comparison, the result of applying the same fitting procedure to the quite spherical lead-glass beads, using the appropriate elastic sphere model in this case, is also shown in Fig. 6 (right half). The values of a_0 and b_0 for the lead-glass beads are 0.99 and 1.06, reflecting the good agreement between spherical scatterer theory and measured data (i.e., within about 5% or better) for these particles.

The result for the 200-μm sand particles in Fig. 6 shows that with the rescaling, the spherical scatterer theory provides an acceptable fit to the data. In Fig. 7 this result is repeated in the two plots at the top and further similar results are shown for the other three size fractions of natural sand and quartz particles listed in Table II. The renormalization works well for all sizes in this range. The quartz particle result is especially interesting because these were the most angular of the particles tested (Fig. 3).

The two-parameter size renormalization also results in a good fit between the rescaled measurements and the theoretical curve for lead-glass beads of different size (Fig. 8; see also Table III) and for natural sand of the same size but different sphericity (Fig. 9; see also Table IV). The results for sands of different sphericity are also interesting. In Fig. 9

it can be seen that the discrepancy between the experimental data and the sphere model is somewhat larger for Assen sand than for Twente and Ottawa sand. This is consistent with a 10% higher value of the parameter b_0 for the scaled measurements. It is noted further that the geological sphericity of the Assen sand is lower than for the other two sands (see Table IV), consistent with the SEM results (see Fig. 5). All values of a_0 and b_0, corresponding to the results in Figs. 6–9, are listed in Tables II, III, and IV. Summary plots showing the variations of these parameters with particle size and type are presented in Fig. 10.

Before discussing the results further, it is necessary to make two remarks about the spectra and the analysis procedure. In the first place, it is noted that in the low-frequency part ($ka<0.5$) of most of the spectra, the experimental data have a relatively large uncertainty, because the value of the attenuation is rather low. This is especially the case for the smallest particle sizes. For lead-glass beads there seem to be systematic trends in the low-frequency data, which, although it should be realized that they are overemphasized because of the $1/f$ weighting, are not understood at present. These data (i.e., the points below $ka=0.5$) were left out of the fitting procedure.

In the second place it should be noted that in the above analysis, a single size has been assumed for the theoretical curves. However, the actual particle size fractions are distributions which have a nonzero width. The effect this has on the fitted values of the two parameters a_0 and b_0 was assessed as follows. Most of the spectra were also fitted using a theoretical curve, convolved with a Gaussian or split-Gaussian size distribution, as an approximation to the actual size distribution. The width of the distribution was taken close to the experimental one as determined by the optical diffraction size analysis. Results in Tables II and III (last columns) show that the effect on the values of a_0 and b_0 is relatively small if the standard deviation of the size distribution is not larger than about 30%. Two examples are shown in Fig. 11. It is seen that for the 60-μm sand size fraction, the fit and the parameter values are only slightly different from the single size case [see Fig. 7(f)]. The other example of the lead-glass 40-μm size fraction, shows only slightly different parameter values but an improved quality of the fit due to the smoothing of the sphere resonances in the theoretical curve [cf. the single size case in Fig. 8(d)]. For the three cases of the present study where the width of the size distribution is 40% or more, the approximation of the actual size distribution by a Gaussian one is rather crude. The resulting values of b_0 are larger than for the single size case. For the two hydrocyclone lead-glass size fractions, the size distribution is so strongly asymmetrical, that a split Gaussian distribution was used to generate the theoretical curve. In these cases also the resulting values of a_0 are different from the single size fits, they are smaller. The main conclusion for the purpose of the present paper is that the effect of including the particle size distribution in the analysis certainly does not result in smaller values for b_0. Also it can be concluded that the trend as regards the variation of the two parameters for the different size fractions of sand (and quartz) and lead-glass par-
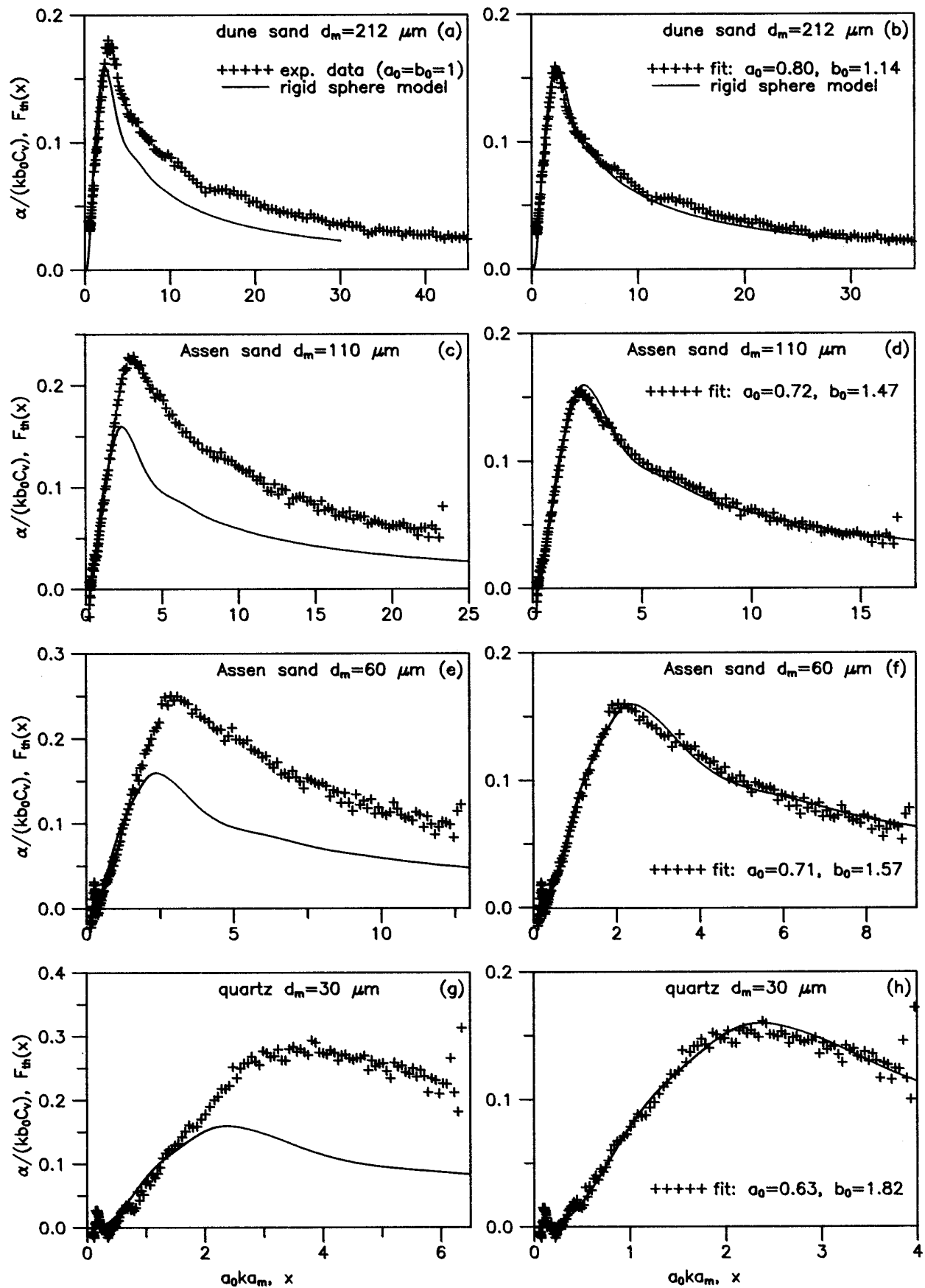
FIG. 7. The left half of the figure shows the measured normalized attenuation spectra of four different size fractions of sand and quartz particles, compared to the rigid movable sphere model for a quartz sphere. The right half of the figure shows the result of scaling the experimental data, using the two-parameter model, to obtain a best fit with the theoretical model.
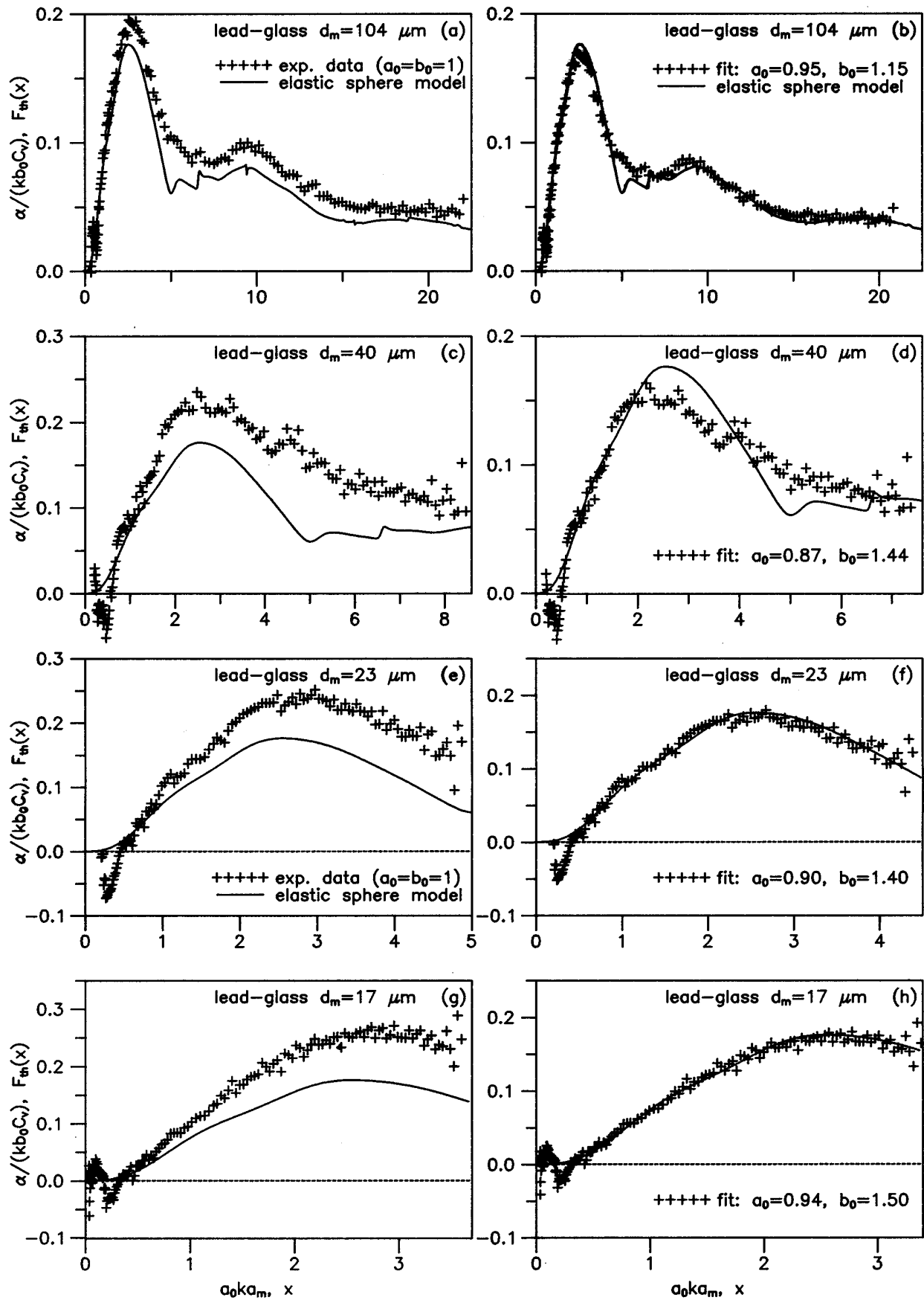
FIG. 8. The left half of the figure shows the measured normalized attenuation spectra of four different size fractions of lead-glass beads, compared to the elastic sphere model for a lead-glass sphere. The right half of the figure shows the result of scaling the experimental data, using the two-parameter model, to obtain a best fit with the theoretical model.
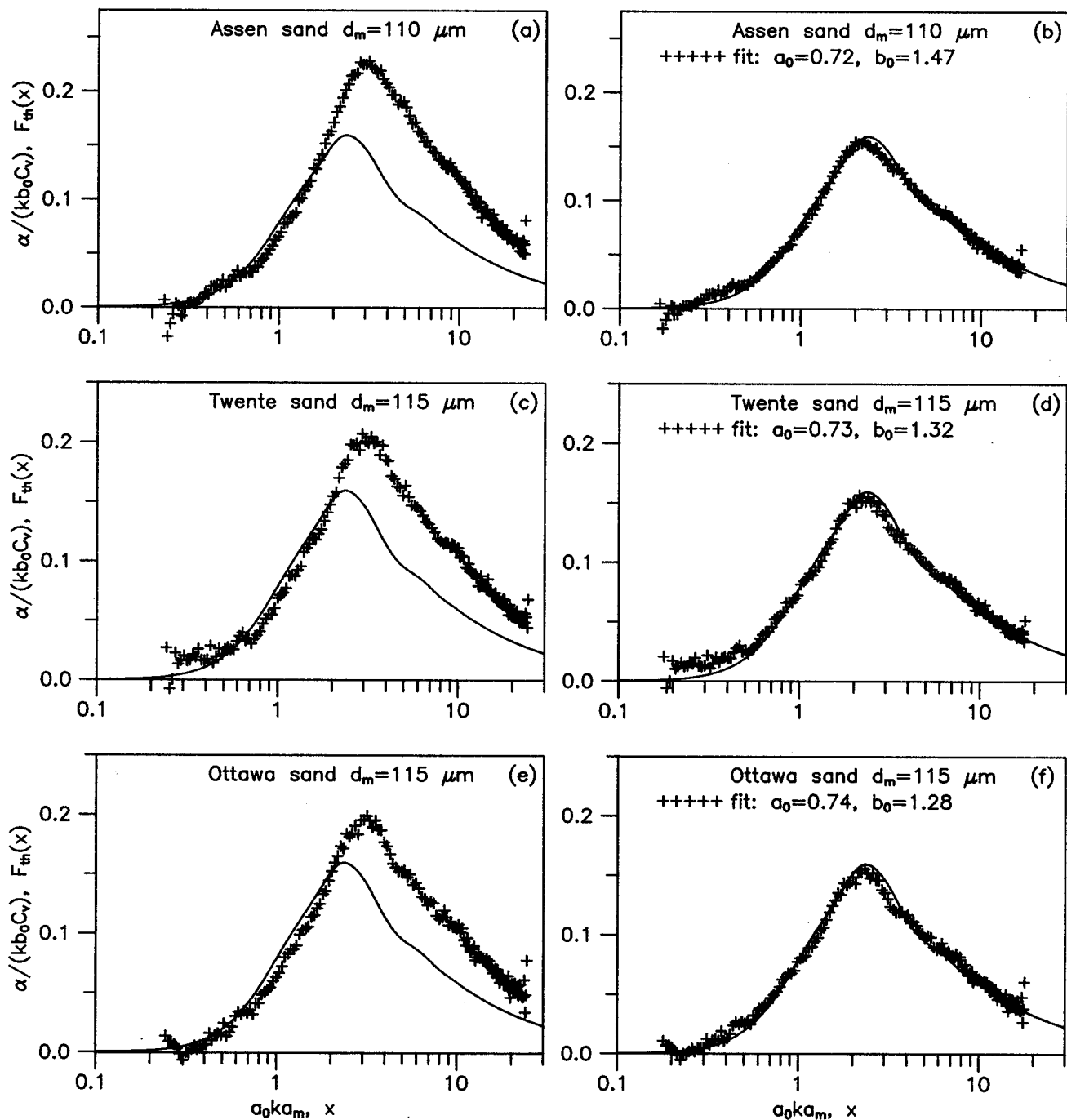
FIG. 9. As-measured, as well as scaled normalized attenuation spectra of 90–106-μm sieve size fractions of three different sands, compared with the rigid movable sphere model for a quartz sphere, are presented in the same way as in the two preceding figures. It is noted that the discrepancy between the experimental data and the sphere model is larger for Assen sand than for Twente and Ottawa sand and this is reflected in the values of the parameter b_0 for the scaled measurements. This difference correlates well with the geological sphericity of the samples (see Table IV).

ticles remains the same, whether or not the size distribution is taken into account.

III. DISCUSSION

A. Physical meaning of b_0

The parameter b_0, as defined by Eq. (9), is determined by the ratio of the two effective volumes V_G and V_p so it is seen that b_0 actually is a measure for the departure from a spherical shape, both in terms of (effective) volume and pro-

jected area. The question is then how the relatively large values of the attenuation and thus of the parameter b_0 for the sand and quartz particles (Figs. 7, 9, and 10; Tables II, IV) can be interpreted in terms of what is known about the shape of the particles. Here the results will be discussed in terms of projected area, since this is most directly related to medium- and high-frequency acoustic cross sections, which were the quantities actually measured.

It follows directly from Eq. (9) that $b_0^{2/3}$ is equal to the ratio of the average projected area of the particle to that of an
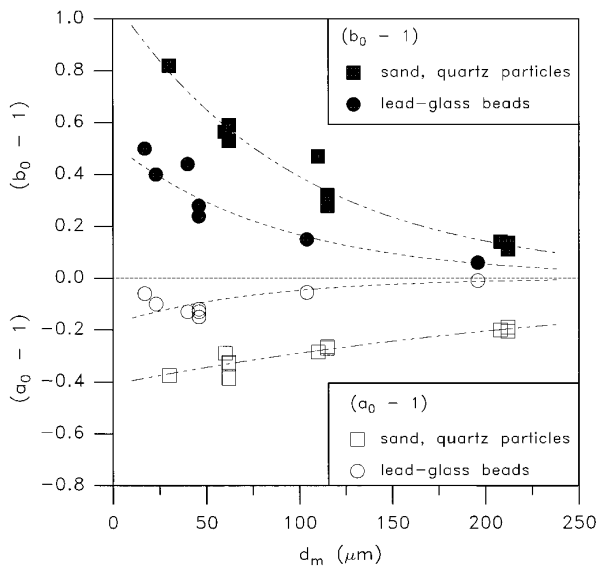
FIG. 10. Summary of all values of the two scaling parameters a_0 and b_0 for the size fractions of sand and quartz particles and lead-glass beads of the present study (see also Tables II, III, and IV).

equal volume sphere. This ratio can be calculated easily for a given convex shape, using a well-known theorem (see, e.g., van der Hulst[31]). The theorem states that the average geometrical cross section of a convex particle with random ori-entation is one-fourth of its surface area. This provides the possibility to compare the experimental values of b_0 with those for some well-known simple polyhedral shapes, which is done in Fig. 12. The interesting point is that the acoustic shape classification for the natural sand and ground quartz particles, provided by this figure, is consistent with the quali-tative shape characterization by scanning electron micros-copy as described in Sec. I. Therefore it appears that the observed values of b_0 are consistent with a shape effect.

For the lead-glass beads, the magnitude of the attenua-tion and so the value of the parameter b_0 is definitely larger than it should be for a perfect sphere, at least for all sizes below 100 μm (Figs. 6, 7, and 10; Table III). The lead-glass beads were chosen to provide an experimental sphere refer-ence for the irregular sediment particles in the same size range. This choice was based on the good agreement be-tween theory and experiment obtained for the 200- and 100-μm size fractions in a previous study.[21] Therefore the present results for the lead-glass beads are unexpected. How-ever, some suggestions can be made to explain the results. First it is noted that the possibility to explain the lead-glass results as a shape effect, in the same way as the sand and quartz particle results, can be ruled out on the basis of the SEM micrographs (Fig. 4). These show that the majority of the lead-glass beads have a very spherical shape and there-fore there will be no increase of the average projected area.
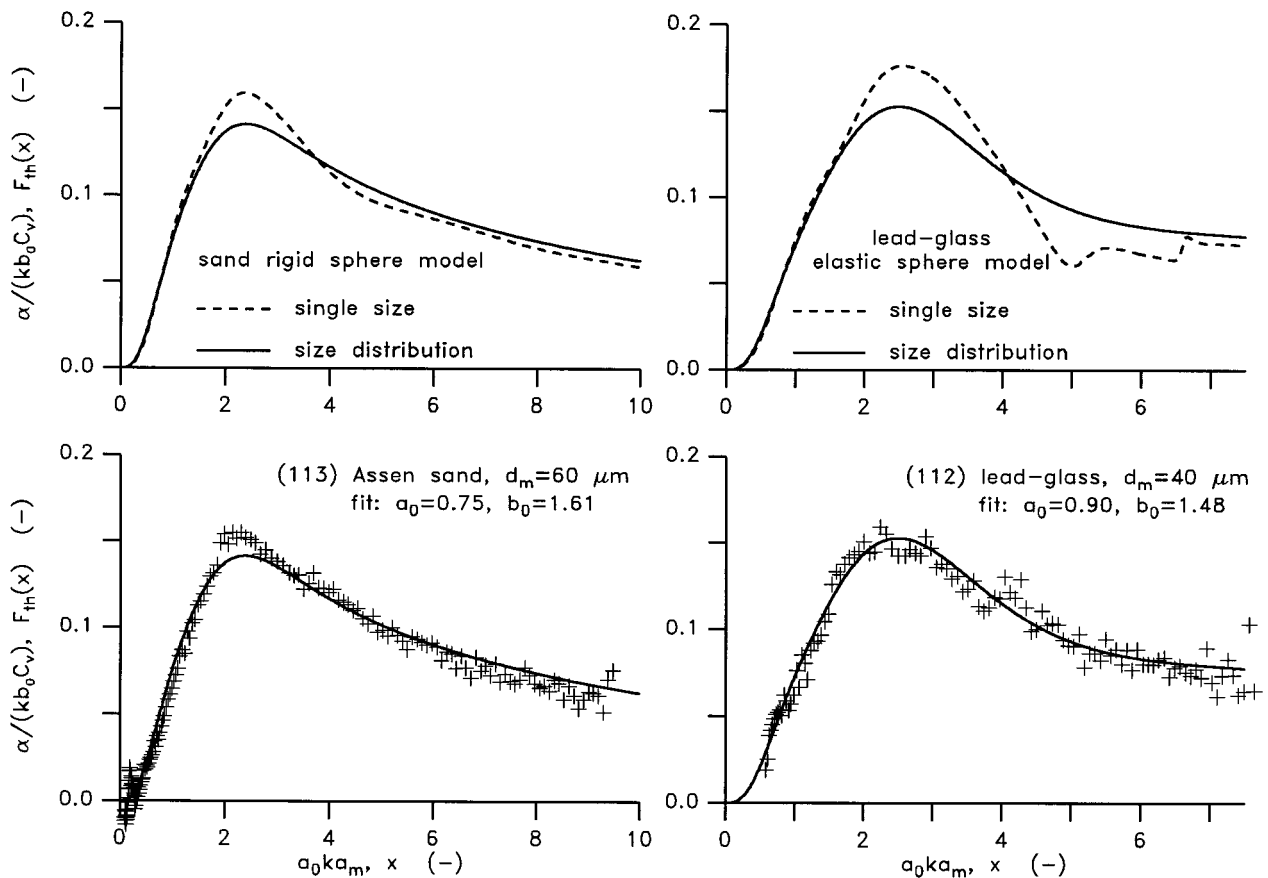
FIG. 11. Example of the effect of taking into account a distribution of sizes, for a Gaussian size distribution with a relative standard deviation of 0.3, which was used to obtain the smoothed (solid) curves for the rigid movable sand or quartz sphere model and the elastic lead-glass sphere model (upper plots). Clearly this also has an effect on the two-parameter scaling procedure (lower plots), where especially the quality of the fit is improved for the case of lead-glass beads (more results are given in Tables II and III).

One suggested explanation is that the extra scattering is due to surface and volume defects (see Sec. I C). However, apart from the observed trend that the number of defects is larger for the smaller size fractions, this contribution cannot be quantified any further without making some assumptions. If it is assumed, for example, that interior holes are responsible for the 12% to 13% lower densities of the smallest lead-glass bead size fractions, this would require a single cavity with a diameter of about half of the diameter of the lead-glass bead. Such cavities would certainly increase the scattering by the beads. Indeed, a strong enhancement of the (backscattering) form function in the ka range between 2 and 10 was found theoretically for hollow metallic spheres in water by Hickling[32] and the enhancement was found to increase with cavity size.

In the present context, it should also be noted that the extra scattering may have been overestimated by 10%–15% for the two smallest size fractions. The reason is the following. In the analysis given up to this point, the same bulk value of the density has been assumed for all size fractions. This means that the lower measured densities for the two smallest size fractions have been ignored. The lower values of the density may be attributed to the presence of interior holes, or alternatively, to a slightly lower lead content of the beads (ignoring the effect of the holes on the density). Actually, whatever their origin, the lower measured densities should be used in the two-parameter model analysis of the attenuation. This leads to proportionally lower values of b_0 [see Eq. (11)]: 1.23 and 1.29 instead of, respectively, 1.40 and 1.49 given in Table III for the two smallest size fractions. These lower values of b_0 are included in Fig. 12 (only).

B. The meaning of a_0

The parameter a_0, as defined by Eq. (8), is equal to the ratio of the acoustic cross-sectional area size a_G and the chosen reference size a_m. The reference size was obtained by a standard optical diffraction instrument, which was used as a black box. It is not the purpose of this paper to discuss discrepancies between the size estimates obtained by the different measurement methods in detail. However, some remarks can be made to indicate that the results are within reasonable limits. The results are visualized in Fig. 10. For the sand and quartz particles the values of a_0 mean that the acoustic size a_G is 20%–40% smaller than the optical diffraction size a_m. For the lead-glass beads, the acoustic size a_G is systematically somewhat lower, up to 15%, than the optical diffraction size estimate. These results are not unreasonable or unexpected. For the spherical lead-glass beads one expects the best agreement, since the optical diffraction instrument is calibrated for spherical particles, as are all standard particle sizing instruments. For the irregularly shaped sediment particles, the optical diffraction instrument will measure some effective size. Here it is relevant to note that Syvitski et al. have found that for a range of sand sediment samples, the size estimates obtained by an electrozone particle sizer were 17%–29% less than those obtained by an optical diffraction instrument.[33] The electrozone instrument

should provide an equal volume size estimate. From the present acoustic attenuation spectra, the equal volume radius can be derived from Eq. (9) as $a_G/b_0^{1/3}$, and using Eq. (8) to obtain a_G. The resulting values are 22%–49% less than the optical diffraction sizes for the sand and quartz particles. Therefore the difference between the present acoustic and optical diffraction size estimates is qualitatively consistent with the finding of Syvitski et al.

It is emphasized that a reference size is not required to determine the acoustic size a_G from a measured attenuation spectrum. Instead of taking some measured size estimate for a_m, one can take any (reasonable) initial guess and then determine a_0 by the fitting procedure to obtain a_G. The values of b_0 do not therefore depend upon a_m and the departure from the expected value of unity (of b_0) for the lead-glass beads, discussed in the previous section, cannot be attributed to errors in the optical diffraction sizes. Thus a_0 is a scaling parameter with no intrinsic physical meaning with respect to particle shape.

C. Comparison with Flammer's data

In a previous study by Sheng and Hay, the measurements made by Flammer of attenuation in aqueous suspensions of sand in the frequency range of 2.5–25 MHz were compared to spherical scatterer theory.[34,18] The result of this comparison is shown in the left half of Fig. 13. It can be seen that as in the present investigation, the measured attenuation is greater than that predicted by rigid sphere theory and that there is also a slight shift to higher values of ka of the position of the main peak. These effects are also larger for the smaller sizes. The size estimates used for the ka scale were the midpoints d_s of the sieve intervals, so the measured radius a_m, as used in the above formal description, is equal to
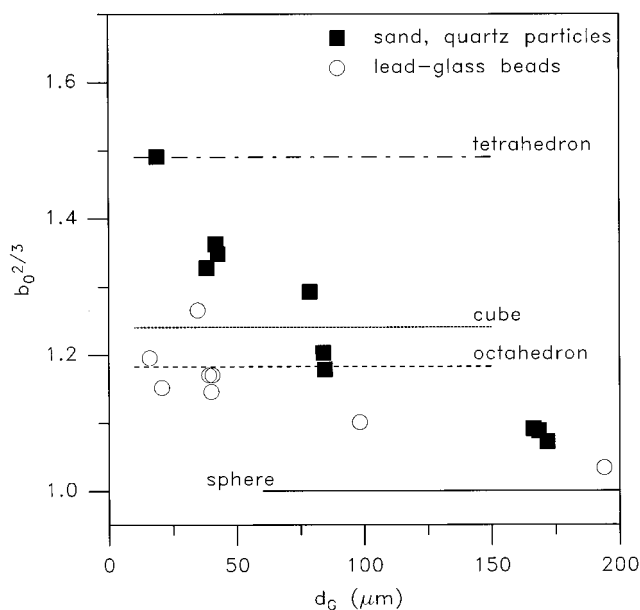
FIG. 12. The values of the parameter $b_0^{2/3}$, which correspond to the acoustically determined projected area relative to that of an equal volume sphere [cf. Eq. (9)], are summarized here as a function of the acoustically determined average size d_G. They are compared with the corresponding values for some simple polyhedral shapes, as further explained in the text.
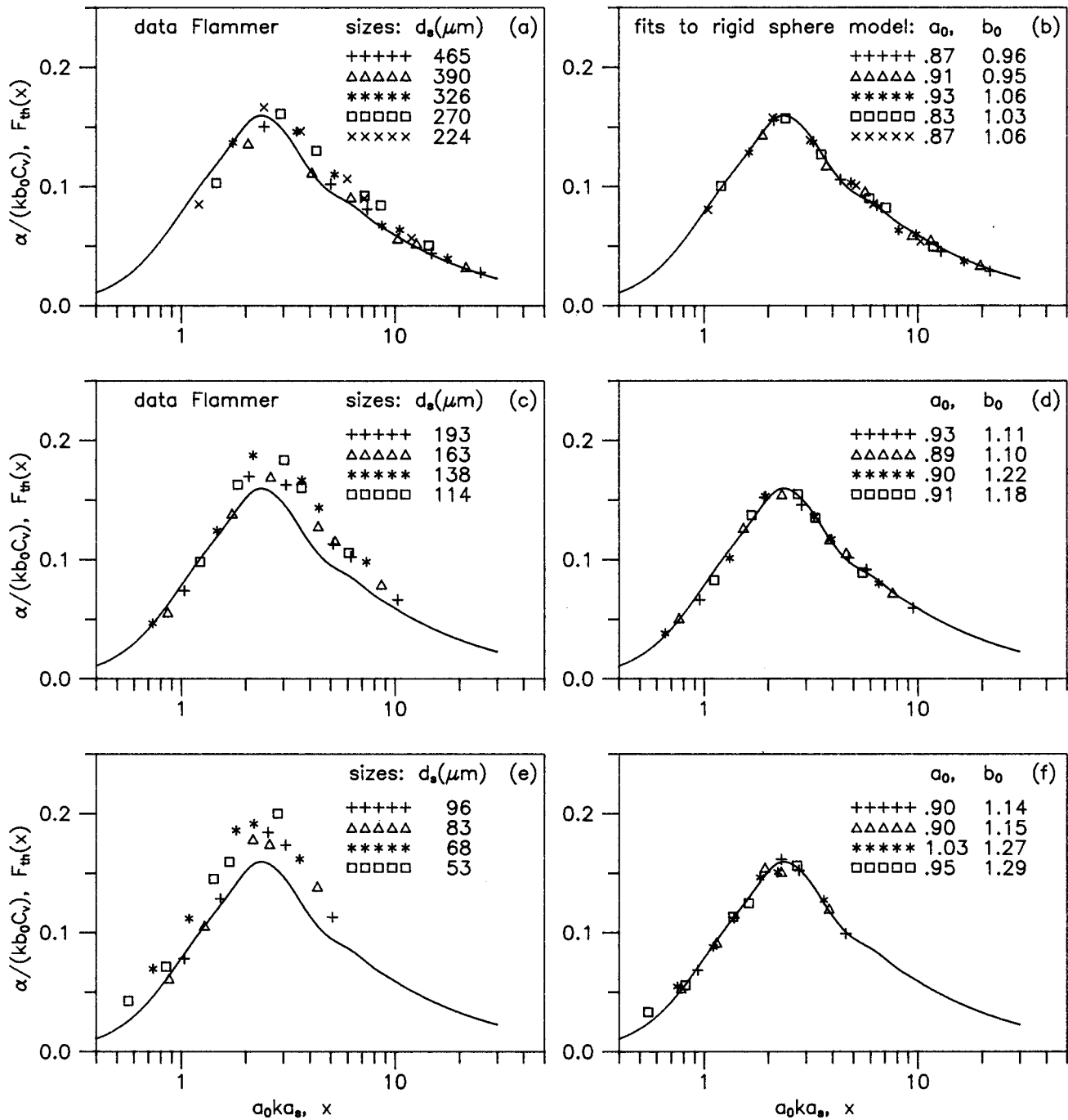
FIG. 13. Experimental data obtained by Flammer on 13 sieve size fractions made up from Missouri River and blasting sand (Ref. 18, Fig. 3, p. A16), compared with the theoretical model for a rigid movable quartz sphere. The data are presented in the same way as in Figs. 7–9 for the present study, with the exception that the average sieve size is used to normalize the acoustic wave number. Note that the average sieve diameters d_s are given with the as-measured spectra in the left half of the figure, and the parameter values a_0 and b_0 with the fitted measurements in the right half of the figure.

$d_s/2$ in the present case. The results of the two-parameter model analysis are given in the right half of Fig. 13, showing that also for these earlier measurements much better fits to the spherical scatterer theory are obtained, although there is quite a lot of scatter in Flammer's data and the spectra consist of a small number of points (maximum 6). It is consistent with the above remarks that the scaling parameters for Flammer's data show the same trends as those for the present study (see Fig. 14).

On the basis of Flammer's measurements, Sheng and Hay proposed that a modified version of Johnson's high-pass model could be used to represent the scattering cross section of natural sand particles,[35] in order to represent the higher measured attenuation compared to the computed values for a rigid sphere. This formulation has been used by a number of investigators to compute the scattering attenuation when inverting backscatter profiles to obtain profiles of suspended sand concentration.[14–16,26] However, in view of the present
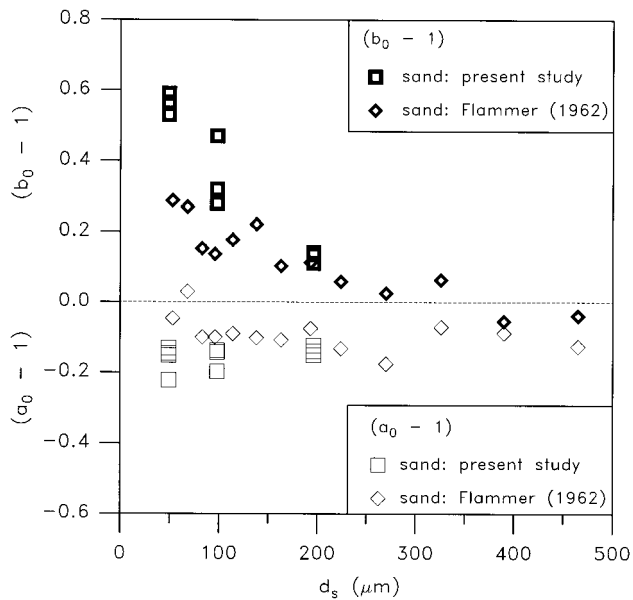
FIG. 14. Comparison of the two scaling parameters for Flammer's data (Fig. 13), with those for the sand sieve size fractions of the present study. Note that also for the latter data, the value of a_0 given is related to the average sieve size, i.e., $a_0 = d_G / d_s$ [see Eq. (8)].

results, it is clear that the rigid sphere model can provide a very good fit to the data if the effects of irregular shape are taken into account. This result is appealing because the two-parameter spherical scatterer model attempts to include the relevant physics, whereas the high-pass model is purely *ad hoc*.

D. Viscous and thermal absorption

Thus far the measured attenuation has been attributed to scattering alone. There is also the possibility that viscous and thermal absorption may have contributed to the measurements. This possibility is briefly considered here.

Thermal absorption by natural sediment particles has been considered elsewhere.[36] It was found that viscous absorption was more important, in the long-wavelength region at least. It can also be argued that thermal losses remain small at shorter wavelengths. This argument is based on the theoretical and experimental work of Allegra and Hawley.[37] They have shown that the thermal losses peak at a value of the ratio of the particle radius to the thermal boundary layer thickness of the order of one. This ratio is much larger than one for all particles of the present study.

Viscous absorption is expected to be important mainly at long wavelengths, as first demonstrated by Urick.[38] An estimate of the contribution of viscous attenuation to the measured attenuation can also be made for medium and high frequencies, using the spherical scatterer theory.[39,40] The results of such calculations made for lead-glass spheres of the different average sizes of the present study can be summarized as follows. The contribution of the viscous attenuation for values of ka between 1 and 10 around the position of the maximum in the spectra of $\alpha/(kb_0 C_\nu)$ (cf. Fig. 6), is about 5% of the attenuation due to scattering alone for the smallest lead-glass particles ($d_{50} = 17\ \mu$m, see Table III) and de-

creases with increasing particle size to about 1.2% for the largest size (i.e., $d_{50} = 196\ \mu$m). Thus if the viscous attenuation were to be included in the theoretical model used for the two-parameter model analysis, the resulting values of b_0 would be lower by maximum 5% for the smallest (lead-glass bead) size.

The above-mentioned results for lead-glass beads also provide an estimate of the importance of viscous absorption in the experiments with sand and quartz particles, as the grain densities of these materials are not markedly different from that of lead-glass. Thus it is estimated that for the sand and quartz particles, the maximum viscous attenuation would be 3% for the quartz particles and this would lead to a correspondingly lower value of b_0. This estimate is rather crude, however, as one would expect the viscous absorption to depend on the total surface area of the particles, and that this would be significantly larger for irregular particles than for spheres of a given equivalent size. Therefore it was felt necessary to test the importance of viscosity experimentally. This was done by measuring the attenuation spectrum for a 45–53-μm-diam sieve size fraction of sand particles as well as of lead-glass beads at three different temperatures: 10, 20, and 30 °C. The viscosity of water changes by 50% over this temperature range, and one would expect to observe a change in the attenuation spectrum if viscous absorption were important. However, no significant changes were observed above $x = 1$, as is indicated by the values of a_0 and b_0 obtained from these measurements (see Tables II and III).

In view of the above discussion, it is concluded that, by neglecting the viscous attenuation, the values of b_0 have been overestimated a little, however, by a maximum 5% only. It may also be concluded from the calculations of the viscous attenuation, that for particle sizes of about 20 μm and smaller it is necessary to include viscous effects in an accurate analysis of attenuation spectra at these frequencies.

IV. CONCLUSIONS

It has been shown that measurements of attenuation spectra in turbulent suspensions of irregular sand and silt-sized particles can be brought into agreement with rigid spherical scatterer theory, but only by invoking two different equivalent sizes. The use of two sizes is appealing on physical grounds. One size is the diameter of a circle with equal projected (i.e., geometric) area, which is the natural scale for the total scattering cross section in the high-frequency limit. This size is used to scale the acoustic wave number, the abscissa of the attenuation spectrum. The second equivalent size is the diameter of the equal volume sphere, which is the natural scale for converting the volume concentration of particles to their number density. This size is used to scale the magnitude of the attenuation, which is the ordinate of the attenuation spectrum. In actual fact, the rescaling factor for the ordinate is a combination of the two equivalent sizes, since the projected area of the particle also affects the magnitude of the attenuation. Because the choice was made to emphasize the region in the vicinity of $ka = 2$, the vertical scaling factor $(1/b_0)$ appears here as the ratio of the actual particle volume to the volume based upon the size of the equal area sphere.

The two-parameter spherical scatterer model provides good agreement not only with the attenuation measurements for suspensions of natural sand and ground quartz presented here, but also with the much earlier measurements in sand suspensions made by Flammer.[18] The attenuation appears to be enhanced, i.e., $b_0 > 1$, in all cases studied. The estimates of b_0 are larger than expected for the smaller-sized lead-glass beads, although still less than the values for irregularly shaped particles in the same size range. It is suggested that inhomogeneities in the composition of the smaller beads (most probably interior defects, inferred from low grain densities and directly observed by scanning electron microscope) are responsible for the observed extra scattering. Unfortunately, therefore, the results obtained for the nominally spherical lead-glass beads do not provide direct support for the above interpretation in terms of a shape effect only, nor do they contradict it. However, direct support for the existence and approximately correct magnitude of the shape effect as found in the present study, is provided by numerical calculations of the orientation average total scattering cross section for a number of irregularly shaped bodies, of which some preliminary results have been published[41] and a comparison with experimental data is underway.[42]

The present results bring into focus the role of grain shape in the propagation of sound through suspensions of moderately irregular particles, natural and otherwise. The effects of grain shape on quantitative estimates of the concentration of suspended particles in marine environments using acoustic remote sensing techniques have largely been ignored. Present results are cautionary, suggesting that shape effects are potentially important, especially in the fine sand and silt-sized range, at the MHz frequencies commonly used in acoustic backscatter systems. This might be seen as further complicating the problem of inverting acoustic backscatter measurements to obtain particle concentrations. The more positive view is that the relatively simple theory for a spherical scatterer provides such good agreement with the measurements. This good agreement implies that spherical scatterer theory can be used in inversion algorithms, provided account is taken of the need for the two equivalent sizes.

ACKNOWLEDGMENTS

[1] R. A. Young, J. T. Merrill, T. L. Clarke, and J. R. Proni, ''Acoustic profiling of suspended sediments in the marine bottom boundary layer,'' Geophys. Res. Lett. **9**, 175–178 (1982).
[2] A. E. Hay, ''On the remote acoustic detection of suspended sediment at long wavelengths,'' J. Geophys. Res. **88**, 7525–7542 (1983).
[3] F. R. Hess and K. W. Bedford, ''Acoustic backscatter system (ABSS): the instrument and some preliminary results,'' Mar. Geol. **66**, 357–380 (1985).
[4] D. M. Hanes and D. A. Huntley, ''Continuous measurements of suspended sand concentration in a wave-dominated nearshore environment,'' Cont. Shelf Res. **6**, 585–596 (1986).
[5] A. S. Schaafsma and W. J. G. J. der Kinderen, ''Ultrasonic instruments for the continuous measurement of suspended sand transport,'' in *Proceedings of the IAHR Symposium on Measuring Techniques in Hydraulic Research*, edited by A. C. E. Wessels (Balkema, Rotterdam, 1986), pp. 125–136.
[6] A. S. Schaafsma and A. J. Wolthuis, ''Frequency dependence of the interaction of ultrasound with suspended sediment particles,'' in *Progress in Underwater Acoustics*, edited by H. M. Merklinger (Plenum, New York, 1986), pp. 153–160.
[7] D. V. Holliday, ''Acoustic determination of suspended particle size spectra,'' Coastal Sediments '87, 260–271 (1987).
[8] K. R. Dyer and R. L. Soulsby, ''Sand transport on the continental shelf,'' Annu. Rev. Fluid Mech. **20**, 295–324 (1988).
[9] D. M. Hanes, C. E. Vincent, D. A. Huntley, and T. L. Clarke, ''Acoustic measurements of suspended sand concentration in the C^2S^2 experiment at Stanhope Lane, Prince Edward Island,'' Mar. Geol. **81**, 185–186 (1988).
[10] J. F. Lynch and Y. C. Agrawal, ''A model-dependent method for inverting vertical profiles of scattering to obtain particle size spectra in boundary layers,'' Mar. Geol. **99**, 387–401 (1991).
[11] J. F. Lynch, T. F. Gross, B. H. Brumley, and R. A. Filyo, ''Sediment concentration profiling in HEBBLE using a 1-MHz acoustic backscatter system,'' Mar. Geol. **99**, 361–385 (1991).
[12] C. E. Vincent and M. O. Green, ''Field measurements of the suspended sand concentration profiles and fluxes and of the resuspension coefficient γ_0 over a rippled bed,'' J. Geophys. Res. **95**, 11591–11601 (1990).
[13] 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).
[14] 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).
[15] P. D. Thorne, P. J. Hardcastle, and R. L. Soulsby, ''Analysis of acoustic measurements of suspended sediments,'' J. Geophys. Res. **98**, 899–910 (1993).
[16] A. E. Hay and J. Sheng, ''Vertical profiles of suspended sand concentration and size from multifrequency acoustic backscatter,'' J. Geophys. Res. **97**, 15661–15677 (1992).
[17] W. G. Bos, ''A comparison of two Doppler current profilers,'' IEEE J. Ocean Eng. **16**, 374–381 (1991).
[18] G. H. Flammer, ''Ultrasonic measurement of suspended sediment,'' Geological Survey Bulletin No. 1141-A (US Government Printing Office, Washington, DC, 1962).
[19] V. K. Varadan, V. N. Bringi, V. V. Varadan, and Y. Ma, ''Coherent attenuation of acoustic waves by pair-correlated random distributions of scatterers with uniform and Gaussian distributions,'' J. Acoust. Soc. Am. **73**, 1941–1947 (1983).
[20] A. S. Schaafsma, ''Acoustic measurement of concentration and size of sediment particle suspensions,'' in *Proceedings Ultrasonics International '89* (Butterworth, Washington, DC, 1989), pp. 388–393.
[21] A. E. Hay and A. S. Schaafsma, ''Resonance scattering in suspensions,'' J. Acoust. Soc. Am. **85**, 1124–1138 (1989).
[22] A. S. Schaafsma, ''Empirical 2-parameter model for acoustic attenuation spectra—Development of silt measuring methods,'' Delft Hydraulics Progress Report B260, 1992, 25 pp.
[23] A. S. Schaafsma, ''Acoustic attenuation spectroscopy of suspended sediments,'' in *2nd European Conference on Underwater Acoustics*, edited by L. Bjørnø (European Commission, Luxembourg, 1994), Vol. 2, pp. 863–868.
[24] A. S. Schaafsma, ''Particles of irregular shape, scattering and multicomponent mixtures—Development of silt measuring methods,'' Delft Hydraulics Progress Report F37, 1991, 47 pp.
[25] A. E. Hay, ''Sound scattering from a particle-laden turbulent jet,'' J. Acoust. Soc. Am. **90**, 2055–2074 (1991).
[26] 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).
[27] C. He and A. E. Hay, ''Broadband measurements of the acoustic back-

scatter cross section of sand particles in suspension,'' J. Acoust. Soc. Am. **94**, 2247–2254 (1993).

[28] F. H. Fisher and V. P. Simmons, ''Sound absorption in sea water,'' J. Acoust. Soc. Am. **62**, 558–562 (1977).

[29] J. M. M. Pinkerton, ''A pulse method for the measurement of ultrasonic absorption in liquids: Results for water,'' Nature (London) **160**, 128–129 (1947).

[30] R. E. Carver, *Procedures in Sedimentary Petrology* (Wiley, New York, 1971).

[31] H. C. van de Hulst, *Light Scattering by Small Particles* (Dover, New York, 1981), p. 110.

[32] R. Hickling, ''Analysis of echoes from a hollow metallic sphere in water,'' J. Acoust. Soc. Am. **36**, 1124–1137 (1964).

[33] J. P. M. Syvitski, K. W. G. Leblanc, and K. W. Asprey, in *Principles, Methods, and Application of Particle Size Analysis*, edited by James P. M. Syvitski (Cambridge U.P., Cambridge, England, 1991), Chap. 13.

[34] J. Sheng and A. E. Hay, ''An examination of the spherical scatterer approximation in aqueous suspensions of sand,'' J. Acoust. Soc. Am. **83**, 598–610 (1988).

[35] R. K. Johnson, ''Sound scattering from a fluid sphere revisited,'' J. Acoust. Soc. Am. **61**, 375–377 (1977).

[36] A. E. Hay and R. W. Burling, ''On sound scattering and attenuation in suspensions, with marine applications,'' J. Acoust. Soc. Am. **72**, 950–959 (1982).

[37] J. R. Allegra and S. A. Hawley, ''Attenuation of sound in suspensions and emulsions: theory and experiments,'' J. Acoust. Soc. Am. **51**, 1545–1564 (1972).

[38] R. J. Urick, ''The absorption of sound in suspensions of irregular particles,'' J. Acoust. Soc. Am. **20**, 283–289 (1948).

[39] A. E. Hay and D. G. Mercer, ''On the theory of sound scattering and viscous absorption in aqueous suspensions at medium and short wavelengths,'' J. Acoust. Soc. Am. **78**, 1761–1771 (1985).

[40] A. E. Hay and D. G. Mercer, ''A note on the viscous attenuation of sound in suspensions,'' J. Acoust. Soc. Am. **85**, 2215–2216 (1989).

[41] T. Mazoyer and D. Guyomar, ''Sediment scattering modelling by boundary elements,'' in *2nd European Conference on Underwater Acoustics*, edited by L. Bjørnø (European Commission, Luxembourg, 1994), Vol. 2, pp. 911–916.

[42] A. S. Schaafsma, A. M. Lafort, T. Mazoyer, and D. Guyomar, ''Characterization of suspended sediment by acoustic techniques, Part I: Theoretical and experimental validation of attenuation spectroscopy,'' Acta Acust. (to be published).

Optimal determination of the material symmetry axes and associated elasticity tensor from ultrasonic velocity data

Christophe Aristégui and Stéphane Baste
Université Bordeaux I, Laboratoire de Mécanique Physique, U.R.A. C.N.R.S. No. 867,
351, Cours de la Libération, 33405-Talence Cedex, France

A simultaneous identification of the angular parallax locating the higher symmetry coordinate system and the associated optimal stiffness tensor from wave speed measurements of obliquely ultrasonic bulk waves in an arbitrarily oriented coordinate system is presented. The property used in classifying a material with regard to its elastic symmetry is the existence and the number of planes of reflective or mirror symmetry. That leads to considering the problem of determining the symmetry class and the directions of the elements of symmetry. To consider the uncertainties of the experimental data, the wave speed measurements are only used to determine the symmetry frames and the optimal stiffness tensor. The proposed inverse propagation algorithm consists of minimizing a functional where the unknowns are the elasticity constants and the Euler angles between the geometric coordinate system and the frame of higher symmetry. Stability of the used least-square algorithm to the initial guesses and to the noise in the wave speed measurements is shown by using simulated data for materials with the most general anisotropy. The applicability of the method is demonstrated using experimental data for arbitrarily oriented but known composite materials. © *1997 Acoustical Society of America.* [S0001-4966(97)05109-6]

PACS numbers: 43.35.Cg, 43.35.Zc [HEB]

INTRODUCTION

Ultrasonic wave speed measurements constitute the most widely used method for the determination of the elasticity constants of anisotropic solids. Usually, these methods require the knowledge of the material symmetry axes. In principle the coincidence between the symmetry axes and the geometric axes of the sample is assumed. However, some discordance between these two coordinate systems may appear in many cases; a possible experimental error of the sample positioning, cutting defects, the strata's stacking defects in industrial composite material, annular rings of wood, microcracks with a predominant orientation that does not coincide with the symmetry axes, etc. That leads to considering the problem of determining the symmetry class and the directions of the elements of symmetry from the knowledge of the general elasticity tensor in an arbitrary coordinate system. In a previous paper,[1] we have demonstrated that the stiffnesses of materials that possess any class of symmetry (including the most general anisotropy) can be extracted from the experimentally measured velocities of obliquely incident ultrasonic waves from water onto a plate. The proposed process for identifying the elasticity tensor takes no account of the symmetry of the acquisition planes. So, this approach is applicable to materials that have a single (monoclinic symmetry) or even no (triclinic symmetry) plane of symmetry.

Even if the media exhibit the most general anisotropy, the intrinsic material symmetry can be greater in one of its symmetry coordinate systems. This can be emphasized by the fourth-rank tensor rotation laws[2] which transform, for example, 21 nonindependent moduli relative to an arbitrary coordinate system into 9 independent moduli relative to a principal coordinate system. The orientation of this higher symmetry frame with respect to the observation coordinate system may be specified by the angular parallax (Euler angles).

A set of measured phase velocities in various directions in a medium contains information both on the moduli and on the location of these directions with respect to the symmetry axes. This information can be used to identify the material symmetry if the stiffnesses are known, or even to determine both the symmetry and the stiffnesses. Privileged directions in a medium, as the preferential alignment of the short fibers in reinforced composites or as the microfacture orientation in rocks, can also be recovered from ultrasonic velocities by using the orientation distribution function.[3,4]

From an ultrasonic contact method[5] for the moduli measurement, Cowin and Mehrabadi have considered[6] the problem of determining the material symmetry of an anisotropic linearly elastic solid. It was shown that each of the ten distinct elastic symmetries can be characterized uniquely by the number and orientation of the planes of symmetry it possesses. So equivalent sets of necessary and sufficient conditions[6–9] upon the components of the fourth-rank modulus tensor for a given direction to be normal to a plane of symmetry were established. From a physical point of view, these directions must be both a specific axis[7] and a specific direction.[10,11] Alternative methods were proposed by minimizing the deviation from symmetry models[12] or by considering three invariants of the Christoffel matrix.[13] All these approaches do not consider the uncertainties of the elasticity constants. To remove this disadvantage, a double-iterative numerical scheme using only the wave speed measurements was proposed for the ''simultaneous'' recovery of the

moduli of a hexagonal specimen and the only ''out-of-plane'' Euler angle.[14]

The problem considered in this paper is to use only the wave speed measurements in an arbitrarily selected coordinate system to determine the symmetry frames for an unknown elastic material. Identification of the material symmetry is carried out by assuming the types of symmetries occurring in materials; triclinic, monoclinic, orthorhombic, tetragonal, hexagonal, and isotropic symmetries, and by determining the more probable symmetry. The monoclinic symmetry is defined by a monoclinic frame whose one base vector is the normal to a plane of symmetry. The recognition of the orthorhombic symmetry is equivalent to locating an elasticity principal coordinate system whose three base vectors are normal to a plane of symmetry. The higher symmetries are identified by estimating the deviation between the stiffness tensor reconstructed in the elasticity principal coordinate system and the one satisfying the chosen model of symmetry.

Simultaneous identification of the angular parallax locating the higher symmetry coordinate system and the associated optimal stiffness tensor are performed from ultrasonic measurements in the initial coordinate system by using nonlinear least-squares optimization fitting. To improve the convergence of the process, an intermediate functional depending jointly on the experimental data and on the angular unknowns is introduced. A first assessment of those unknowns is obtained using the average eigenvectors of the Voigt tensor and of the dilatational modulus.[6] The nonlinear dependence of the angles on the identified stiffnesses leads naturally to error propagation in calculations, and so, to a rough estimate of the angles. Once the approximate angular parallax are determined, the stiffness transformation laws are used to obtain the numerical values of the elasticity tensor in the symmetry coordinate system. These results are then selected to be the initial guesses of the method. The reliability of the inversion procedure will be established by recognizing *a posteriori*, from variances and covariances delivered by the optimization procedure, which Euler angles and which stiffnesses have been most accurately recovered and which less so.

Numerical examples that simulate in the best way a typical experiment are used to investigate the influence of the data scatter and of the initial guesses on the accuracy of the material symmetry reconstruction. Then the recovering method is applied to experimental data. First, the effect of a defective control of the sample positioning is emphasized. Due to errors induced by a possible wrong recognition of the material symmetry, reconstruction of the stiffness tensor must be carried out by making the fewest assumptions possible on the coincidence between the observation frame and the symmetry elements. The study of materials that possess one known plane of symmetry is treated in Sec. III. The location of the elasticity principal coordinate system is then equivalent to the reconstruction of a single Euler angle and of nine elasticity constants. The case of the materials that possess no known plane of symmetry is presented in Sec. IV. The orientation of the elasticity principal coordinate system with respect to the observation coordinate system may be

specified by the three Euler angles. Identification of the monoclinic symmetry is presented as a simplification of the orthorhombic case since the number of angles (two) necessary for the location of the normal to a plane of symmetry is lower than the one (three) used for an elasticity principal coordinate system. The applicability of the method is demonstrated for arbitrarily oriented but known composite materials.

I. RECOVERY OF THE ELASTICITY TENSOR OF GENERAL ANISOTROPIC MATERIALS

The phase velocities of an elastic plane wave are the solution of the well-known Christoffel equation:[2]

$$|\Gamma_{ij} - \rho V^2 \delta_{ij}| = 0, \tag{1}$$

with $\Gamma_{ij} = C_{ijkl} n_k n_l$ $(i,j,k,l=1,2,3)$ and where ρ is the density, $\mathbf{n}$ is the unit vector in the wave propagation direction, V is the phase velocity of ultrasonic waves in the medium, C_{ijkl} is the elasticity constants of an anisotropic medium, and δ_{ij} is the Kronecker symbol. By inverting Eq. (1), the material properties (C_{ijkl}) can be determined from phase velocity measurements in a suitable set of propagation directions.

Using the abbreviated subscript notation, the elasticity tensor (C_{ijkl}) is expressed by the 21 components of a symmetric (6×6) matrix (C_{IJ}):

$$(C_{IJ}) = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ & & C_{33} & C_{34} & C_{35} & C_{36} \\ & & & C_{44} & C_{45} & C_{46} \\ & \text{Sym.} & & & C_{55} & C_{56} \\ & & & & & C_{66} \end{bmatrix}. \tag{2}$$

According to the degree of material anisotropy, the number of unknowns, that is, the number of independent elasticity constants, varies from 2 to 21 for the isotropic to the triclinic symmetry. Materials with monoclinic symmetry have a single plane of symmetry. When this single plane is the plane $(\mathbf{x}_2, \mathbf{x}_3)$, the stiffness tensor (C_{IJ}) takes the form

$$(C_{IJ}) = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & & \\ & C_{22} & C_{23} & C_{24} & & \\ & & C_{33} & C_{34} & 0 & \\ & & & C_{44} & & \\ & \text{Sym.} & & & C_{55} & C_{56} \\ & & & & & C_{66} \end{bmatrix}. \tag{3}$$

When the three planes $(\mathbf{x}_1, \mathbf{x}_2)$, $(\mathbf{x}_1, \mathbf{x}_3)$, and $(\mathbf{x}_2, \mathbf{x}_3)$ are planes of symmetry and are mutually orthogonal, the material presents at least an orthorhombic symmetry. Its elasticity tensor (C_{IJ}) has nine nonzero components:

FIG. 1. Schematic diagram of the simple-transmission experiment. (a) $R = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ is the geometric coordinate system associated with the sample. The incident plane $(\mathbf{x}_1, \Psi_i)$ is defined by the azimuthal angle Ψ_i. (b) Ultrasonic wave propagation in the incident plane $(\mathbf{x}_1, \Psi_i)$. The incident θ_i and refraction θ_r angles are reported for positive values.

$$(C_{IJ}) = \begin{bmatrix} C_{11} & C_{12} & C_{13} & & & \\ & C_{22} & C_{23} & & 0 & \\ & & C_{33} & & & \\ & & & C_{44} & & \\ \text{Sym.} & & & & C_{55} & \\ & & & & & C_{66} \end{bmatrix}. \qquad (4)$$

A. Phase velocity measurements

The geometric coordinate system $R = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ is chosen in such a way that axis $\mathbf{x}_1$ corresponds to the normal to the interface [Fig. 1(a)]. Wave speed measurements are performed by using ultrasonic pulses which are transmitted through a plate immersed in water [Fig. 1(b)]. Ultrasound was generated and sensed by a pair of piezoelectric transducers. They are supported by two goniometers which allow one to select any propagation direction $\mathbf{n}_i$ of incident wave [Fig. 1(b)]. For an oblique incident angle ($\theta_i \neq 0°$), the longitudinal incident pulse is refracted at the liquid–solid interface and generates one, two, or three waves that propagate with different velocities along the directions $\mathbf{n}_m$, $m = 1$, 2, or 3. One quasilongitudinal (QL), one fast (QT1), and one slow (QT2) quasitransverse waves can generally be excited in each incident plane. If this plane coincides with a plane of symmetry, only one quasitransverse (QT) wave will be generated in the solid. One of these waves has been reported in Fig. 1(b) and noted $\mathbf{n}$ ($\mathbf{n} = \mathbf{n}_m$, with $m = 1$, 2, or 3).

A signal processing of the received pulses is performed by using a reference signal propagating in water without the sample. As demonstrated for this experimental device,[15,16] the time delay caused by the presence of the anisotropic plate

FIG. 2. Choice of a nonprincipal observation coordinate system R' that does not coincide with the geometric coordinate system R of a unidirectional composite. R' is obtained from a 5° rotation of $R = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ about the $\mathbf{x}_1$ axis ($\phi = 5°$). The axis $\mathbf{x}_3$ is along to the fiber direction.

is equal to the time δt for an acoustic wave propagating with the phase velocity inside the material. Among the signal processings used for the time-delay measurements, the Hilbert transform is from many points of view the most appropriate determination of the stiffnesses, independently from the attenuation mechanisms.[17,18] Using the Hilbert transform the velocity at normal incidence is measured from the time-of-flight difference between the through-sample and the fluid paths. Taking the transmitted signal at normal incidence as reference[19–21] improves the accuracy of the δt measurement. The advantage of such a method is especially important when dispersion effects are present.

Phase velocities V_p at incident angle θ_i, imposed by the goniometers, are collected in several incident planes $(\mathbf{x}_1, \Psi_i)$, where Ψ_i is the azimuthal angle [Fig. 1(a)]. The refraction angles $\theta_r = (\mathbf{x}_1, \mathbf{n}_m)$, $m = 1$, 2, or 3, of the phase velocities are directly related to the angle of incidence through Snell–Descartes' laws.[2]

B. Stiffness tensor reconstruction

Due to the experimental errors, every measured phase velocity V_p is approximately the solution of the nonlinear cubic equation $f_p(v_p(\mathbf{n}), C_{IJ})$, the left-hand side of the Christoffel equation, Eq. (1). Identification of the stiffness tensor from wave speed measurements in suited acquisition planes $\cup$(data planes) is then performed by minimizing the functional

$$F(C_{IJ}) = \sum_{V_p \in \cup(\text{data planes})} \{f_p(V_p(\mathbf{n}), C_{IJ})\}^2, \qquad (5)$$

built from the overdetermined system of the functions $f_p(v_p(\mathbf{n}), C_{IJ})$. For a successful reconstruction of the stiffness tensor, the experimental velocity set should be sensitive to the elasticity constant set to be determined and must express the specific symmetry to be reconstructed. For an orthorhombic medium, the three data planes,[22]

$$\cup(\text{data planes}) = (\mathbf{x}_1, \mathbf{x}_2) \cup (\mathbf{x}_1, \mathbf{x}_3) \cup (\mathbf{x}_1, 45°), \qquad (6)$$

are sufficient to identify the nine components of the stiffness tensor (C_{IJ}). Determination of the additional elasticity constants, 4 and 12 for a medium with respectively monoclinic

C. Aristégui and S. Baste: Material symmetry axes   1505

TABLE I. Results (in GPa) of the elasticity constant (C'_{IJ}) reconstruction for a carbon epoxy composite using simulated ultrasonic velocity data (a) in the planes $(\mathbf{x}'_1,\mathbf{x}'_2)$, $(\mathbf{x}'_1,\mathbf{x}'_3)$, $(\mathbf{x}'_1,(45°)')$ and $(\mathbf{x}'_1,(135°)')$, assuming that the observation frame $R'=(\mathbf{x}'_1,\mathbf{x}'_2,\mathbf{x}'_3)$ does not coincide with the elasticity principal coordinate system R^p; and (b) in the planes $(\mathbf{x}'_1,\mathbf{x}'_2)$, $(\mathbf{x}'_1,\mathbf{x}'_3)$ and $(\mathbf{x}'_1,(45°)')$, assuming that the observation frame R' coincides with the frame R^p. The numbers in the parentheses indicate the confidence interval (in GPa) associated with each elasticity constant.

Data type	C'_{11}	C'_{12}	C'_{13}	C'_{14}	C'_{22}	C'_{23}	C'_{24}	C'_{33}	C'_{34}	C'_{44}	C'_{55}	C'_{56}	C'_{66}
Original data	12.15	5.43	5.46	0.0	12.42	6.46	0.68	128.6	9.57	7.6	5.98	0.22	3.5
(a)	12.15 (0.00)	5.43 (0.00)	5.46 (0.00)	0.0 (0.0)	12.42 (0.00)	6.46 (0.00)	0.68 (0.00)	128.6 (0.0)	9.57 (0.00)	7.6 (0.0)	5.98 (0.00)	0.22 (0.00)	3.5 (0.0)
(b)	12.1 (0.1)	5.5 (0.2)	5.6 (0.3)	0	12.3 (0.2)	13.2 (1.2)	0	139 (5)	0	8.4 (0.6)	6.2 (0.1)	0	3.6 (0.1)

and triclinic symmetry, requires one to consider additional data planes. Contrary to the orthorhombic case, the $(\mathbf{x}_1,\mathbf{45°})$ and $(\mathbf{x}_1,\mathbf{135°})$ data planes of a medium, possessing a single symmetry plane $(\mathbf{x}_2,\mathbf{x}_3)$, are different. Consequently, comparison of the data collected respectively in these planes allows one to differentiate the orthorhombic symmetry from the monoclinic symmetry. Introduction[1] of the additional data plane $(\mathbf{x}_1,\mathbf{135°})$,

$$\cup(\text{data planes})=(\mathbf{x}_1,\mathbf{x}_2)\cup(\mathbf{x}_1,\mathbf{x}_3)\cup(\mathbf{x}_1,\mathbf{45°})\cup(\mathbf{x}_1,\mathbf{135°}),\tag{7}$$

is sufficient to recover the stiffness tensor (3). The identification of the most general stiffness tensor, Eq. (2), is carried out from the wave speed measurements in the four data planes, Eq. (7). Due to the nonsymmetry plane $(\mathbf{x}_2,\mathbf{x}_3)$, the data measurements are performed for positive and negative incident angles.

C. Wrong recognition of the symmetry material

To emphasize the effect of a defective control of the sample positioning, a unidirectional composite is characterized in an observation frame $R'=(\mathbf{x}'_1,\mathbf{x}'_2,\mathbf{x}'_3)$ different from an elasticity principal coordinate system R^p, which coincides *a priori* with the geometric frame R (Fig. 2). A small deviation between R' and R^p has been chosen to reproduce the possible experimental error. The density of the studied sample of carbon epoxy composite is 1.56 g/cm^3 and the thickness is 3.45 mm. The stiffness tensor (C^p_{IJ}) relative to R^p is[1]

$$(C^p_{IJ})=\begin{pmatrix} 12.15\pm0.03 & 5.4\pm0.1 & 5.5\pm0.2 & & & \\ & 12.3\pm0.2 & 5.7\pm0.8 & & 0 & \\ & & 130\pm3 & & & \\ & & & 6.8\pm0.1 & & \\ & \text{Sym.} & & & 6.00\pm0.04 & \\ & & & & & 3.48\pm0.04 \end{pmatrix} \ (\text{GPa}).\tag{8}$$

Here R' is obtained from a 5° rotation of R about the $\mathbf{x}_1$ axis. Using the stiffness tensor transformation, the 13 equivalent elasticity constants in R' are determined from (C^p_{IJ}) [Eq. (8)]. From these 13 ''original data'' (given in Table I), a set of phase velocities was calculated for selected propagation directions. The moduli identification is then performed in R' from this synthetic velocity set.

Since R and R' are very close, the angular ranges experimentally accessible in each plane $(\mathbf{x}'_1,\mathbf{x}'_2)$, $(\mathbf{x}'_1,\mathbf{x}'_3)$, $(\mathbf{x}'_1,(45°)')$, and $(\mathbf{x}'_1,(135°)')$ are close to the ones of the planes $(\mathbf{x}_1,\mathbf{x}_2)$, $(\mathbf{x}_1,\mathbf{x}_3)$, and $(\mathbf{x}_1,\mathbf{45°})$ [$=(\mathbf{x}_1,\mathbf{135°})$]. The range of refraction angles for each of these data planes is determined from the experimental velocities[22] measured in this carbon epoxy sample. In the plane $(\mathbf{x}_1,\mathbf{x}_2)$, the range of

the refraction angle for the QL mode is [0°; 71°] and is [25°; 60°] for the QT mode. The number of the simulated data is 23 for the QL mode and 12 for the QT mode. In the plane $(\mathbf{x}_1,\mathbf{x}_3)$, the range of the refraction angle is [0°;12°] for the QL mode and [12.5°;70°] for the QT mode (with 13 and 23 synthetic data). The ranges of the refraction angle for the QL, QT1, and QT2 modes propagating in the plane $(\mathbf{x}_1,\mathbf{45°})$ are respectively [0°;18°], [15°; 70°], and [22°; 68°] (with 9, 33, and 20 simulated data).

In R' the specimen is ''apparently'' monoclinic. Contrary to the planes $(\mathbf{x}'_1,\mathbf{x}'_2)$ and $(\mathbf{x}'_1,\mathbf{x}'_3)$, the plane $(\mathbf{x}'_2,\mathbf{x}'_3)$ [$=(\mathbf{x}_2,\mathbf{x}_3)$] preserves its symmetric character. The 13 elasticity constants (C'_{IJ}) are identified from the simulated data in the four planes $(\mathbf{x}'_1,\mathbf{x}'_2)$, $(\mathbf{x}'_1,\mathbf{x}'_3)$, $(\mathbf{x}'_1,(45°)')$, and

C. Aristégui and S. Baste: Material symmetry axes

FIG. 3. Slownesses for a unidirectional carbon epoxy composite in the (a) $(\mathbf{x}_1', \mathbf{x}_2')$, (b) $(\mathbf{x}_1', \mathbf{x}_3')$, and (c) $(\mathbf{x}_1', (45°)')$ planes. The points are simulated slownesses without scatter. The lines are slownesses calculated from the reconstructed elasticity constants taking into account the monoclinic (solid line) or orthorhombic (dashed line) symmetry model.

$(\mathbf{x}_1', (\mathbf{135°})')$ [Table I(a)]. As in all the tables, the numbers in parentheses represent the confidence interval associated with each elasticity constant. Since the phase velocities are simulated without scatter, the reconstructed moduli are naturally identical to the original data. Assuming that the observation coordinate system R' coincides with an elasticity principal frame, the nine components of (C_{IJ}') are determined from the three data planes $(\mathbf{x}_1', \mathbf{x}_2')$, $(\mathbf{x}_1', \mathbf{x}_3')$, and $(\mathbf{x}_1', (45°)')$ [Table I(b)]. The insensitivity of the confidence intervals to the systematic errors explains that the identified elasticity constants, with their respective confidence interval, do not coincide with the original data. The identification of the moduli C_{23}', C_{33}', and C_{44}' is highly dependent on the wrong choice of the stiffness tensor form. The slowness curves calculated from the nine moduli determined considering the orthorhombic model illustrate this defective characterization (Fig. 3). The effect of this erroneous recognition of the symmetry is greater for the reconstruction of the QT2 mode in the plane $(\mathbf{x}_1', (45°)')$ and of the QT mode in the planes $(\mathbf{x}_1', \mathbf{x}_2')$ and $(\mathbf{x}_1', \mathbf{x}_3')$.

II. RECOVERY OF THE MATERIAL SYMMETRY

Each elastic symmetry is characterized[6] by the number and orientation of the planes of symmetry it possesses. The monoclinic symmetry is defined by a monoclinic frame, R^m, whose one base vector is the normal to a plane of symmetry. The location of an elasticity principal coordinate system, R^p, whose three base vectors are normal to a plane of symmetry allows one to identify the orthorhombic symmetry. The higher symmetries (tetragonal, hexagonal or isotropic) are identified by estimating the deviation between the stiffness tensor reconstructed in R^p and the one satisfying the chosen model of symmetry.

The orientation of a higher symmetry frame R^s ($= R^m$ or R^p) with respect to the observation coordinate system R may be specified by the Euler angles (ψ, ϕ, θ). These angles relate the base vectors of R and R^s through three successive rotations about different coordinate axes. The coordinates are first rotated clockwise through an angle ψ about the $\mathbf{x}_3$ axis, then through a clockwise angle ϕ about the transformed $\mathbf{x}_1$ axis, and finally rotated clockwise through an angle θ about the transformed $\mathbf{x}_3$ axis. The location of a material symmetry frame R^s is equivalent to determining the Euler angles (ψ, ϕ, θ) which transform from R to R^s. The reconstruction of the angular unknowns, $\delta = (\psi, \phi, \theta)$, must allow us to identify the material symmetry. The base vectors of R^s are then related to the ones of R by the (3×3) rotation matrix $(P_{IJ})_\delta = P_\theta \circ P_\phi \circ P_\psi$. The components (C_{IJ}) and (C_{IJ}^s) represent the same stiffness tensor in two different coordinate systems R and R^s. The relation between these tensors is

$$[C_{IJ}^s] = [M_{IJ}]_\delta [C_{IJ}][M_{IJ}]_\delta^t, \tag{9}$$

where $(M_{IJ})_\delta$ is the (6×6) transformation matrix defined[2] from $(P_{IJ})_\delta$. The subscript $(M_{IJ})_{\delta^{-1}}$ denotes the matrix determined from the inverse matrix of $(P_{IJ})_\delta$, $(P_{IJ})_{\delta^{-1}}$.

A. Research of an elasticity principal coordinate system R^p

When the sample possesses a single known plane of symmetry, for example, the plane $(\mathbf{x}_2, \mathbf{x}_3)$, the single nonzero Euler angle ϕ allows one to locate an elasticity principal frame $R^p = (\mathbf{x}_1^p, \mathbf{x}_2^p, \mathbf{x}_3^p)$ (if it is exists) with respect to the observation coordinate system R. The angular unknowns δ are reduced to (ϕ). On the other hand, if no assumption about material symmetry can be made, the three Euler angles (ψ, ϕ, θ) are necessary to search for R^p.

1. Location of R^p from the moduli identified in the observation frame R

A first estimate of the unknowns δ may be extracted from the stiffness (C_{IJ}). When the medium symmetry is "apparently" monoclinic, the angular parallax ϕ is determined from the 13 nonzero components (C_{IJ}) [Eq. (3)], using the 9 invariants[23] relative to the fourth-rank tensor rotation around the $\mathbf{x}_1$ axis:

$$C_{11}, \quad C_{12}+C_{13}, \quad C_{22}+C_{33}+2C_{23},$$
$$C_{55}+C_{66}, \quad C_{44}-C_{23}, \quad C_{12}C_{13}-C_{14}^2,$$
$$C_{55}C_{66}-C_{56}^2, \quad (C_{33}-C_{22})^2+4(C_{24}+C_{34})^2, \tag{10}$$
$$C_{22}^2+C_{33}^2+2C_{23}^2+4(C_{24}^2+C_{34}^2+C_{44}^2).$$

Assuming the existence of R^p, the nine elasticity constants (C_{IJ}^p) are then calculated. The value of ϕ is extracted from the analytic expressions of the stiffness transformation between (C_{IJ}) and (C_{IJ}^p) [Eq. (9)].

Another approach, also independent of the uncertainties of the stiffnesses (C_{IJ}) but applicable to the components of the general triclinic tensor given in an arbitrary coordinate system, has been presented in Ref. 6. From the conditions[6] on the existence of a symmetry plane, Arts proposes[12] to use one particular result deduced from the two symmetric tensors of rank 2, namely the dilatational modulus (d_{IJ}) and the Voigt tensor (V_{IJ}):

$$V_{IJ}=C_{ikkj}, \quad d_{IJ}=C_{ijkk}. \tag{11}$$

For materials possessing at least three mutually orthogonal planes of symmetry, the eigenvectors of (V_{IJ}) and (d_{IJ}) are identical and aligned with the crystallographic directions. For monoclinic media, these two tensors have a single eigenvector in common (normal to the plane of symmetry). For the triclinic case, the two sets have no eigenvector in common. However, due to the experimental errors in the measurements, the eigenvectors of (V_{IJ}) and (d_{IJ}) do not exactly line up. Since the tensors (11) are symmetric, both (V_{IJ}) and (d_{IJ}) each have three mutually orthogonal eigenvectors $(\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3)$ and $(\mathbf{d}_1,\mathbf{d}_2,\mathbf{d}_3)$, respectively. Thus the normals to a plane of symmetry are contained in a solid angle around the average directions, V_I^a ($I=1,2,3$), between the three closest eigenvectors of (V_{IJ}) and (d_{IJ}) (Fig. 4). A fairly good estimate of the angular parallax can be obtained if the moduli (C_{IJ}) are accurately identified. In that case, the solid angles are small.

However, the tensors (V_{IJ}) and (d_{IJ}) being built from linear combinations of the stiffnesses (C_{IJ}) [Eq. (11)], theirs eigenvectors depend nonlinearly on these moduli (C_{IJ}). Therefore, the error propagation in these calculations affects strongly the assessment of the angular parallax δ.

2. Location of R^p from the wave speed measurements in the observation frame R

A method, taking directly into account the wave speed measurements, is presented. The existence of R^p is assumed. Let (C_{IJ}) be the stiffness tensor identified in R. A functional $A(\delta)$ is built that is minimal for a value δ^p locating R^p with
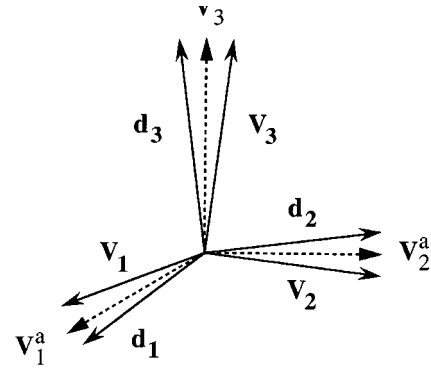
FIG. 4. Determination of the mean vectors $\mathbf{V}_I^a$ ($I=1$ to 3) which are candidates for a normal to a plane of symmetry. $(\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3)$ are the eigenvectors of the Voigt tensor (V_{IJ}) and $(\mathbf{d}_1,\mathbf{d}_2,\mathbf{d}_3)$ are the ones of the dilatational modulus (d_{IJ}).

respect to R. For any value δ_0 of the unknowns δ, the coordinate system $R_{\delta_0}=(\mathbf{x}_1^{\delta_0},\mathbf{x}_2^{\delta_0},\mathbf{x}_3^{\delta_0})$ is defined by

$$\mathbf{x}_I^{\delta_0}=(P_{IJ})_{\delta_0}\mathbf{x}_J, \tag{12}$$

with $I,J=1$ to 3, and where $(P_{IJ})_{\delta_0}$ is the (3×3) rotation matrix relating the frames R and R_{δ_0}. Using Eq. (9), the stiffness tensor $(C_{IJ})_{\delta_0}$ relative to R_{δ_0} is calculated from (C_{IJ}) and δ_0:

$$(C_{IJ})_{\delta_0}=(M_{IJ})_{\delta_0}(C_{IJ})(M_{IJ})_{\delta_0}^t. \tag{13}$$

The Christoffel equation is rewritten in R_{δ_0}, assuming that this frame coincides with an elasticity principal coordinate system. The propagation direction $\mathbf{n}$ is expressed in R_{δ_0}:

$$\mathbf{n}(\delta_0)=(P_{IJ})_{\delta_0}^{-1}\mathbf{n}. \tag{14}$$

The symmetry character of the searched frame is specified by the form of $(C_{IJ})_{\delta_0}$. The nine components of $(C_{IJ})_{\delta_0}$ relative to the orthorhombic symmetry, Eq. (4), are reported in the Christoffel equation. Using Eq. (5)'s notations, the functional can be written as

$$A(\delta_0)=\sum_{V_p\in(\cup\text{data planes in }R)}\{f_p(V_p(\mathbf{n}(\delta_0)),(C_{IJ})_{\delta_0}^{\text{model}})\}^2, \tag{15}$$

with

$$(C_{IJ})_{\delta_0}^{\text{model}}=\text{model}[(C_{IJ})_{\delta_0}],$$

and where the function *model* $[(C_{IJ})_{\delta_0}]$ applies the form (4) on the tensor $(C_{IJ})_{\delta_0}$. If an elasticity principal coordinate system exists, the unknowns $\delta^\mathbf{p}$ are determined by solving

$$\delta^\mathbf{p}=\min_{\delta_0} A(\delta_0). \tag{16}$$

The coefficients (C_{IJ}) appear implicitly in $A(\delta)$ through the tensor $(C_{IJ})_{\delta_0}$ [Eq. (13)], without considering the uncertain-

ties $I(C_{IJ})$. However, due to the taking into account of the experimental data, the solution of the problem (16) is naturally better than the estimate presented in Sec. II A 1.

Let N_A be the number of the unknowns of the problem (16). When the plane $(\mathbf{x}_2, \mathbf{x}_3)$ [Fig. 1(a)] is a plane of symmetry, there exists a single solution. When the ''apparent'' material symmetry is triclinic, N_A is equal to 3. The angular parallax assessment deduced from the (V_{IJ}) and (d_{IJ}) eigenvectors is selected to be the initial guess of the problem (16). The accuracy of the optimization results is estimated calculating the confidence interval[24] $I(\delta)$ associated with each component of the identified vector $\delta^{\mathbf{p}}$. Using the stiffness transformation laws, the numerical values $(C_{IJ}^p)^{\text{computed}}$ of the stiffness tensor relative to $R^{\mathbf{p}}$ are obtained from (C_{IJ}) and δ^p without taking into account $I(\delta)$.

B. Simultaneous recovery of an elasticity principal coordinate system and of the optimal elasticity tensor

The optimal determination of the stiffness tensor (C_{IJ}^p) and of the angular triplet $\delta^{\mathbf{p}}$ locating an elasticity principal coordinate system with respect to the observation frame R is carried out from the wave speed measurements in R. The results of the minimization of the functional $A(\delta)$ and the moduli $(C_{IJ}^p)^{\text{computed}}$ are chosen as the initial guesses of the reconstruction algorithm.

For any value δ_0 of the unknowns δ, nine nonzero components of the elasticity tensor $(C_{IJ})_{\delta_0}$ relative to the frame R_{δ_0} are selected. In R, this tensor becomes

$$(C_{IJ})_R = (M_{IJ})_{\delta_0}^{-1} (C_{IJ})_{\delta_0} (M_{IJ})_{\delta_0}^{t-1}. \tag{17}$$

According to the ''apparent'' material symmetry of the frame which coincides with R, the number of elasticity constants $(C_{IJ})_R$ varies: respectively 13 and 21 for the monoclinic and triclinic symmetries. This tensor is reported in the Christoffel equation [Eq. (5)], and allows one to define the functional

$$B((C_{IJ})_{\delta_0}, \delta_0) = \sum_{V_p \in (\cup \text{data planes in } R)} \{f_p(V_p(\mathbf{n}), (C_{IJ})_R)\}^2. \tag{18}$$

The moduli (C_{IJ}^p) and the angular triplet $\delta^{\mathbf{p}}$ can be found by minimizing

$$((C_{IJ}^p), \delta^p) = \min_{((C_{IJ})_{\delta_0}, \delta_0)} B((C_{IJ})_{\delta_0}, \delta_0). \tag{19}$$

1. Reconstruction algorithm

The number N_B of unknowns, solutions of the problem (19), depends on the ''apparent'' material symmetry in R. It is equal to 10 when the symmetry is monoclinic ($N_B = N_A + 9$) and 12 (3 angles and 9 C_{IJ}) for the triclinic case. Let $\mathbf{Y}$ be the $(1 \times N_B)$ vector formed by the N_B unknowns. The minimization of the functional $B(\mathbf{Y})[=B((C_{IJ}), \delta)]$ is carried out[25] by using successively the steepest descent method and the generalized Newton method. The iterative schemes of these two successive approaches are identical and can be written as

$$\mathbf{Y}_{k+1} = \mathbf{Y}_k + \tau_k \mathbf{Z}_k, \tag{20}$$

with $\mathbf{Y}_k$, $\mathbf{Z}_k \in \mathfrak{R}^{N_B}$ and $\tau_k \in \mathfrak{R}^+$ ($\tau_k > 0$). The expression of $\mathbf{Z}_k$ with respect to the components of $\mathbf{Y}_k$ varies according to the used method. Classically $\mathbf{Z}_k$ is defined from the $(1 \times N_B)$ gradient vector $\mathbf{g}_k$ and from the symmetric $(N_B \times N_B)$ Hessian matrix H_k of the functional (18) with respect to the N_B unknowns. The steepest descent method is characterized by

$$\mathbf{Z}_k = -g_k / \|g_k\|, \tag{21}$$

whereas the generalized Newton method is built from

$$\mathbf{Z}_k = -H_k^{-1} g_k. \tag{22}$$

The step τ_k is introduced to minimize the functional (18) along the direction $\mathbf{Z}_k$. Here τ_k is the solution of the halfway problem:

$$\tau_k = \min_{\tau \in \mathfrak{R}^+} B^*(\tau) = \min_{\tau \in \mathfrak{R}^+} B(\mathbf{Y}_k + \tau \mathbf{Z}_k). \tag{23}$$

Solving this problem is performed[25] by using the dichotomy method with derivatives and the Newton–Raphson method. However, due to the form of $B(\mathbf{Y})$, the optimal τ_k is only a local minimum.

An estimation of the confidence interval $I(\mathbf{Y})$ associated with each component of $\mathbf{Y}$ can be calculated by means of a statistical analysis of the set of velocity measurements in $\cup$(data planes). The values $I(\mathbf{Y})$, deduced from a linearization of the function $f_p(v_p(\mathbf{n}), (C_{IJ})_R)$ [Eq. (18)], around the exact solution for each measured velocity, are the diagonal terms of the $(N_B \times N_B)$ covariance matrix.[24] This matrix is built from the $(N_B \times N)$ gradient matrix ($N =$ total number of data) and the N error vector components. The gradient matrix contains the gradient of $f_p(v_p(\mathbf{n}), (C_{IJ})_R)$ with respect to the N_B unknowns for any data. The error vector collects the final value of the function for any data. Let $\mathbf{Y}^{(\text{int})}$ be the $(1 \times N_B)$ vector that contains the Euler angles $\delta^{(\text{int})}$ locating (with respect to R) an intrinsic elasticity principal frame and the associated moduli $(C_{IJ})^{(\text{int})}$. On assuming a Student's law, the intrinsic quantities $\mathbf{Y}^{(\text{int})}$ will be such that

$$|\mathbf{Y}_i^{(\text{int})} - \mathbf{Y}_i| < I(\mathbf{Y}_i), \quad i = 1 \text{ to } N_B, \tag{24}$$

with, for example, a 99% probability. Since some assumptions on the statistical distribution of the measured data around the optimized solution are required, that does not provide an exact calculation.

2. Existence of an elasticity principal coordinate system R^p

Solving the problem (19) has been carried assuming the existence of R^p. Using the stiffness transformation laws, the moduli $(C_{IJ})^{\text{computed}}$ relative to the observation frame R can be calculated from the optimal couple $((C_{IJ}^p), \delta^p)$. The good agreement between $(C_{IJ})^{\text{computed}}$ and the stiffness tensor (C_{IJ}) identified from the experimental data validates the solution of the problem (19). The comparison between the stiffness tensors $(C_{IJ})^{\text{computed}}$ and (C_{IJ}) is performed by calculating the deviation[12]

$$\Delta = \sqrt{\frac{(C_{ijkl} - C_{ijkl}^{\text{computed}})^2}{(C_{ijkl})^2}}. \tag{25}$$

The velocities standard deviation[26] $\sigma((C_{IJ})^{\text{computed}}, V_p)$ based on the comparison between the experimental velocity set $\cup$(data planes) and velocities that are computed from the stiffnesses $(C_{IJ})^{\text{computed}}$, quantifies the accuracy of the material symmetry identification.

3. Identification of the class of symmetry

If the frame R^p exists, the material symmetry can be higher than the orthorhombic symmetry. For each considered symmetry: tetragonal, hexagonal, and isotropic, a stiffness tensor (C_{IJ}^{model}) whose components satisfy the relationships relative to these symmetries is built from the moduli (C_{IJ}^p). For example, in the case of hexagonal symmetry when the plane $(\mathbf{x}_1^p, \mathbf{x}_2^p)$ is isotropic, (C_{IJ}^{Hex}) is defined by

$$C_{11}^{\text{Hex}} = C_{22}^{\text{Hex}} = \tfrac{1}{8}(3(C_{11}^p + C_{22}^p) + 2C_{12}^p + 4C_{66}^p),$$

$$C_{33}^{\text{Hex}} = C_{33}^p,$$

$$C_{44}^{\text{Hex}} = C_{55}^{\text{Hex}} = \frac{C_{44}^p + C_{55}^p}{2}, \quad C_{66}^{\text{Hex}} = \frac{C_{11}^{\text{Hex}} - C_{12}^{\text{Hex}}}{2}, \tag{26}$$

$$C_{12}^{\text{Hex}} = \tfrac{1}{8}(C_{11}^p + C_{22}^p + 6C_{12}^p - 4C_{66}^p),$$

$$C_{13}^{\text{Hex}} = C_{23}^{\text{Hex}} = \frac{C_{13}^p + C_{23}^p}{2}.$$

For the others classical symmetries, the relationships between (C_{IJ}^p) and (C_{IJ}^{model}) can be easily established. For each tensor (C_{IJ}^{model}) reconstructed in this way in R^p, the equivalent tensor $(C_{IJ})^{\text{computed}}$ in R is obtained from δ^p. The calculation of the deviation Δ_{model} [Eq. (25)] between $(C_{IJ})^{\text{computed}}$ and the moduli (C_{IJ}) determined from the experimental data allows one to quantify the deviation from the assumed symmetry. The more probable symmetry is then identified.

C. Research of the monoclinic coordinate system R^m

Identification of the monoclinic symmetry is carried out by searching for a monoclinic coordinate system $R^m = (\mathbf{x}_1^m, \mathbf{x}_2^m, \mathbf{x}_3^m)$ in which the form of the stiffness tensor (C_{IJ}^m) satisfies Eq. (3). The normal to the plane of symmetry is then parallel to $\mathbf{x}_1^m$. The R^m identification is equivalent to the reconstruction of this base vector. The two other base vectors $\mathbf{x}_2^m$ and $\mathbf{x}_3^m$ are chosen arbitrarily to form an orthonormal frame. In that case, the two Euler angles ϕ and θ are sufficient to locate R^m with respect to the observation coordinate system R (Fig. 5). The angular unknowns vector δ is reduced to (ϕ, θ).

The process of R^m reconstruction is similar to that described previously for the location of an elasticity principal frame. The determination of the (V_{IJ}) and (d_{IJ}) eigenvectors and the successive minimizations of the functionals $A(\delta)$ [Eq. (15)] and $B(\mathbf{Y})$ [Eq. (18)] allow one to define R^m. The search for a single plane of symmetry implies the selection of one of the three couples formed by the closest eigenvectors of (V_{IJ}) and (d_{IJ}) (Fig. 4). The average direction associated with the vector couple whose angular deviation is minimal is chosen as a first estimate. If the material symmetry is higher than monoclinic, the three average directions $\mathbf{V}_I^a$ $(I = 1,2,3)$ can be selected to be the initial guesses. The functional $A(\delta)$ is written in terms of the two components of δ $(N_A = 2)$. The expression of $A(\delta)$ varies through the function $model$ $[(C_{IJ})_{\delta_0}]$: the tensors (C_{IJ}^m) and $(C_{IJ})_{\delta_0}$ [Eq. (13)] must take the form (3). The periodicity of the two Euler angles ϕ and θ minimizing Eq. (15) is $(\pi/2)$. Due to the symmetric character of R^m, the number of unknowns of the functional $B(\mathbf{Y})$ is now equal to 15 $(N_B = 15)$. The solving algorithm is similar to the one described in Sec. II B. Confidence intervals are associated with the two components of $\delta^{\mathbf{m}}$ and with the 13 elasticity constants (C_{IJ}^m). The estimate of the deviation (25) between the tensors $(C_{IJ})^{\text{computed}}$ [obtained from the optimal solution $((C_{IJ}^m), \delta^{\mathbf{m}})$] and (C_{IJ}) allows one to validate the existence of R^m.

FIG. 5. Location of the monoclinic coordinate system $R^m = (\mathbf{x}_1^m, \mathbf{x}_2^m, \mathbf{x}_3^m)$ with respect to the geometric frame R with the two Euler angles ϕ and θ. $\mathbf{x}_1^m$ is along the normal to the plane of symmetry.

III. APPLICATION TO MATERIALS THAT POSSESS ONE KNOWN PLANE OF SYMMETRY

For composite plates, the plane $(\mathbf{x}_2, \mathbf{x}_3)$ [Fig. 1(a)] is the more natural plane of symmetry. The stiffness tensor (C_{IJ}) takes at least the form (3). If an elasticity principal frame R^p exists, the normal $\mathbf{x}_1$ to the symmetry plane $(\mathbf{x}_2, \mathbf{x}_3)$ is collinear to one of the base vectors of R^p. With no loss in generality, one can take the base vector $\mathbf{x}_1^p$ along the direction of $\mathbf{x}_1$. The search for R^p is then equivalent to the reconstruction of the single Euler angle ϕ associated with the rotation about the $\mathbf{x}_1$ axis $(N_A = 1)$. The periodicity of the angular parallax minimizing the functional $A(\delta)$ is $\pi/2$.

A. Monoclinic media characterization: Numerical simulation

Let R' be an observation coordinate system that is different from the elasticity principal frame to be reconstructed.
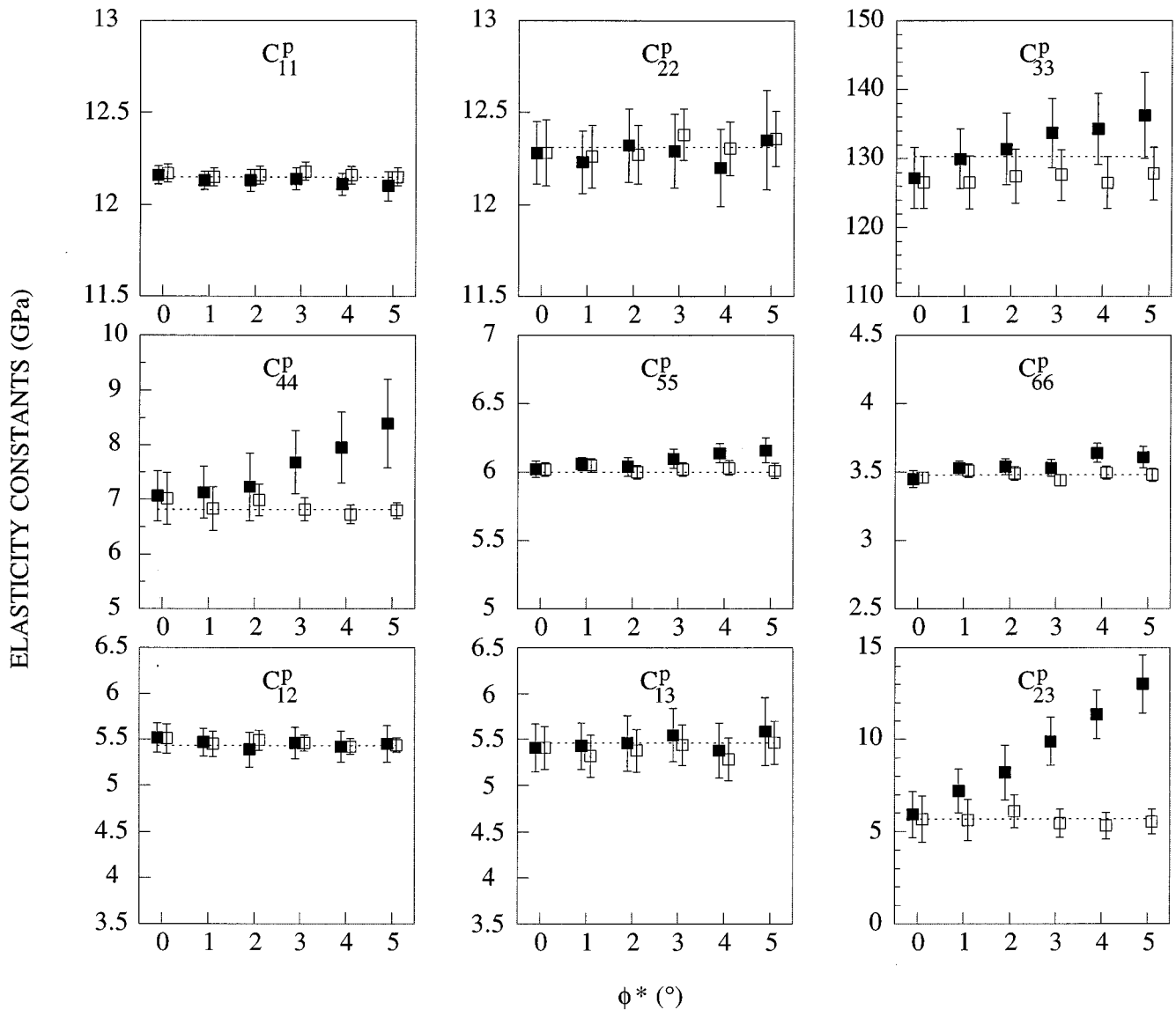
FIG. 6. Reconstruction of the equivalent stiffness tensor in the elasticity principal coordinate system R^p from the simulated data in the observation coordinate systems R' obtained from a ϕ^*-degree rotation of the geometric coordinate system R about the $\mathbf{x}_1$ axis. We assume that the R' frame coincides (solid squares) or does not coincide (empty squares) with the R^p coordinate system. The original values of (C_{IJ}^p), Eq. (8), are shown by dashed lines.

R' is obtained from a ϕ^*-degree rotation of R about the $\mathbf{x}_1$ axis. The symmetry planes $(\mathbf{x}_2', \mathbf{x}_3')$ and $(\mathbf{x}_2, \mathbf{x}_3)$ are superimposed. The reconstructed parallax angular ϕ^p must be equal to $-\phi^*$. The physical properties of the carbon epoxy sample studied in Sec. I C are selected for the simulation. From (C_{IJ}^p), Eq. (8), and ϕ^*, the equivalent stiffness tensor (C_{IJ}') in R' is calculated using the stiffness transformation laws. The 13 components of (C_{IJ}') are used to generate synthetic wave velocity data in the planes $(\mathbf{x}_1', \mathbf{x}_2')$, $(\mathbf{x}_1', \mathbf{x}_3')$, $(\mathbf{x}_1', (45°)')$, and $(\mathbf{x}_1', (135°)')$. The selection of the refraction angles range for each data plane has already been done in Sec. I C. To approach the experimental conditions, it is useful to introduce random scatter. A random error ΔV is added to each computed velocity data V. The velocities are randomized within an interval $\Delta V/V$ of a fixed percentage. A 0.5% level of scatter in the data is sufficient[1,21,27] to reproduce the possible experimental error.

For any value ϕ^* ($\phi^*=0°$, 1°, 2°, 3°, 4°, or 5°), the search for an elasticity principal coordinate system R^p is carried out from the simulated data in the planes $(\mathbf{x}_1', \mathbf{x}_2')$, $(\mathbf{x}_1', \mathbf{x}_3')$, $(\mathbf{x}_1', (45°)')$, and $(\mathbf{x}_1', (135°)')$. The location of R^p is performed by solving the problem (19). Each of the nine moduli (C_{IJ}^p) is represented in Fig. 6 by empty squares. These identified elasticity constants and the original moduli [Eq. (8)] (shown by dashed lines) coincide, to the nearest confidence intervals. The comparison between δ^p and ϕ^* illustrates the accuracy of the R^p location (Fig. 7). The method proposed is sensitive to small angular deviations between the observation frame R' and R^p.

The error due to an erroneous recognition of the material symmetry in R' is emphasized. For each value ϕ^*, we assume that R' coincides with R^p. The stiffness tensor (C_{IJ}') then takes the form (4). The nine moduli (C_{IJ}') identified from the simulated data in the planes $(\mathbf{x}_1', \mathbf{x}_2')$, $(\mathbf{x}_1', \mathbf{x}_3')$, and
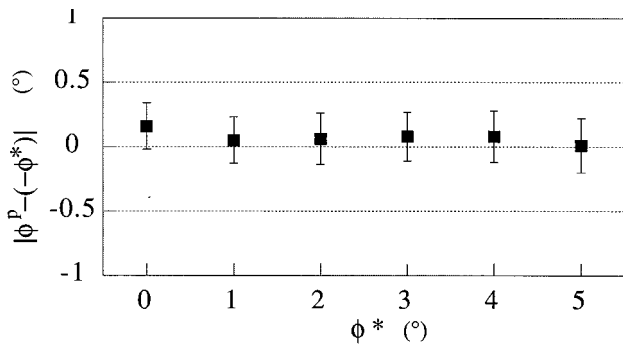
FIG. 7. The angular parallax ϕ^p reconstructed from the simulated data in the observation coordinate systems R' obtained from a ϕ^*-degree rotation of the geometric coordinate system about the $\mathbf{x}_1$ axis.

FIG. 8. Four data planes $(\mathbf{x}_1', \mathbf{x}_2')[=(\mathbf{x}_1, 30°)]$, $(\mathbf{x}_1', \mathbf{x}_3')[=(\mathbf{x}_1, 120°)]$, $(\mathbf{x}_1', (45°)')[=(\mathbf{x}_1, 75°)]$, and $(\mathbf{x}_1', (135°)')[=(\mathbf{x}_1, 165°)]$ for characterization of a unidirectional carbon epoxy composite in the nonprincipal coordinate system $R' = (\mathbf{x}_1', \mathbf{x}_2', \mathbf{x}_3')$.

$(\mathbf{x}_1', (45°)')$ are shown in Fig. 6 by solid squares. For the constants C_{23}^p, C_{33}^p, and C_{44}^p, the quality of determination decreases as the deviation from the intrinsic symmetry of the medium increases. This result confirms the study carried out in Sec. I C. Calculating the engineering constant set from the identified stiffnesses (C_{IJ}'), it can be seen that the identification of the Poisson ratios ν_{13} and ν_{23} and of the shear modulus G_{23} is very defective when ϕ^* increases.

B. Monoclinic media characterization: Experimental results

The unidirectional carbon epoxy composite is now examined in a nonprincipal coordinate system R'. Axis $\mathbf{x}_3$ of the geometric frame R corresponds to the fiber direction. Figure 8 shows R' obtained from a 30° rotation of R about the $\mathbf{x}_1$ axis. Since only the plane $(\mathbf{x}_2', \mathbf{x}_3')$ preserves its symmetric character, the equivalent stiffness tensor (C_{IJ}') in R' takes the form (3). The elasticity constants

$$(C_{IJ}') = \begin{pmatrix} 12.15\pm0.02 & 5.3\pm0.2 & 5.7\pm0.2 & -0.1\pm0.1 & 0 & 0 \\ & 21\pm1 & 24\pm1 & 13.6\pm0.5 & 0 & 0 \\ & & 83\pm2 & 36.9\pm0.7 & 0 & 0 \\ & & & 26.0\pm0.5 & 0 & 0 \\ & \text{Sym.} & & & 5.3\pm0.1 & 1.1\pm0.1 \\ & & & & & 4.2\pm0.1 \end{pmatrix} \text{(GPa)} \qquad (27)$$

are determined from the velocity measurements in the data planes $(\mathbf{x}_1', \mathbf{x}_2')$, $(\mathbf{x}_1', \mathbf{x}_3')$, $(\mathbf{x}_1', (45°)')$, and $(\mathbf{x}_1', (135°)')$. Figure 9 shows the measured data and the slowness curves calculated from the reconstructed elasticity constants [Eq. (27)]. Contrary to the orthorhombic media, the experimental data collected respectively in the data planes $(\mathbf{x}_1, 45°)$ and $(\mathbf{x}_1, 135°)$ are different [Fig. 9(c) and (d)]. The success of such an identification is pointed out by small associated confidence intervals $I(C_{IJ}')$ and by the value of $\sigma((C_{IJ}'), V_p)$ ($=0.4\%$). Reconstruction of 13 elasticity constants does not imply a monoclinic symmetry. It can be higher in an elasticity principal coordinate system.

The R^p identification is carried out from the wave speed measurements in the planes $(\mathbf{x}_1', \mathbf{x}_2')$, $(\mathbf{x}_1', \mathbf{x}_3')$, $(\mathbf{x}_1', (45°)')$, and $(\mathbf{x}_1', (135°)')$. The results of the successive reconstructions of the angular parallax ϕ^p (locating R^p with respect to R') from the nine invariants [Eq. (10)], then from the minimization of the functional $A(\delta)$ [Eq. (15)], are summarized in Table II. For each identified value ϕ^p the stiffness tensor

$(C_{IJ}^p)^{\text{computed}}$ is calculated from (C_{IJ}') using the stiffness transformation laws. The value $\phi^p(=-29.6°)$ and the associated moduli $(C_{IJ}^p)^{\text{computed}}$ are chosen as the initial guess of the $B(\mathbf{Y})$'s minimization. The reconstructed stiffness tensor (C_{IJ}^p) must be absolutely identical to the one given in Eq. (8). These two identifications have been indeed performed from the same experimental data set. The assumption which is too restrictive on the correct knowledge of R^p for the direct determination of the nine elasticity constants [Eq. (8)] can explain the disagreement between, for example, the reconstructions of C_{33}^p (130 GPa in one case, 131 GPa in the other). This is typical of an experimental defective sample positioning. This is also confirmed by the noncoincidence of the identified angular parallax $\phi^p(=-29.6° \pm 0.2°)$ with the simulated angle ($-30°$).

To verify the existence of R^p, the stiffness tensor $(C_{IJ}')^{\text{computed}}$ is determined from the optimal couple $((C_{IJ}^p), \phi^p)$ (Table II):
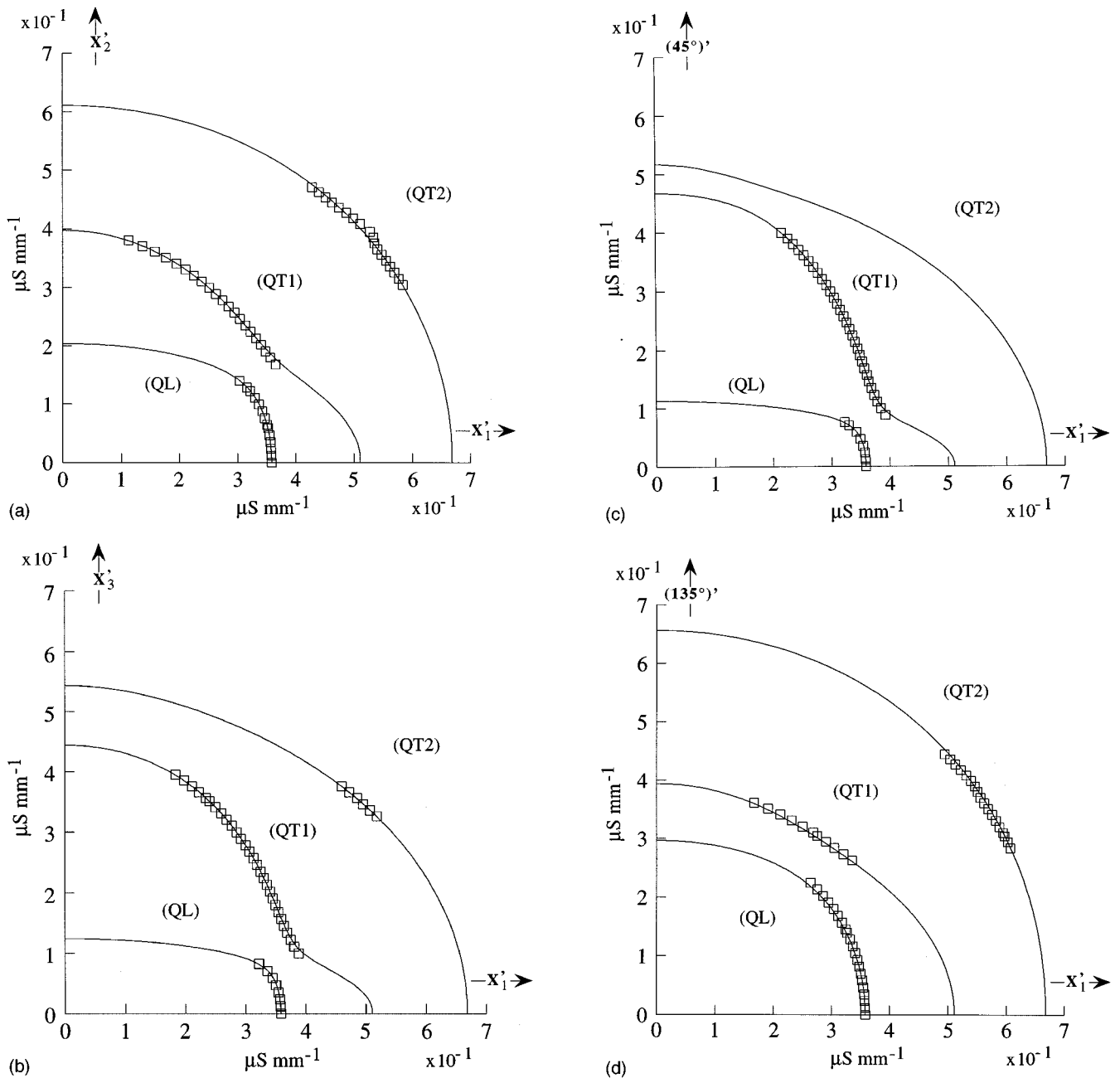
FIG. 9. Slownesses for a unidirectional carbon epoxy composite in the (a) $(\mathbf{x}_1', \mathbf{x}_2')$, (b) $(\mathbf{x}_1', \mathbf{x}_3')$, (c) $(\mathbf{x}_1', (45°)')$, and (d) $(\mathbf{x}_1', (135°)')$ planes. The points are experimentally measured slownesses. The solid lines are slownesses calculated from the reconstructed elasticity constants in the nonprincipal coordinate system $R' = (\mathbf{x}_1', \mathbf{x}_2', \mathbf{x}_3')$.

$$(C_{IJ}')^{\text{computed}} = \begin{pmatrix} 12.15 & 5.45 & 5.46 & 0.01 & 0 & 0 \\ & 21.96 & 24.85 & 13.97 & 0 & 0 \\ & & 82.59 & 36.84 & 0 & 0 \\ & & & 26.06 & 0 & 0 \\ & \text{Sym.} & & & 5.36 & 1.07 \\ & & & & & 4.09 \end{pmatrix} \text{(GPa).}$$

(28)

The standard deviation between the experimental data in the planes $(\mathbf{x}_1', \mathbf{x}_2')$, $(\mathbf{x}_1', \mathbf{x}_3')$, $(\mathbf{x}_1', (45°)')$, and $(\mathbf{x}_1', (135°)')$ and the slowness curves recalculated with the stiffness tensor (28) is $\sigma((C_{IJ}')^{\text{computed}} V_{,p})$ $(=0.36\%)$. The deviation Δ $(=1.1\%)$ [Eq. (25)] between this computed tensor and the moduli (C_{IJ}') directly identified in R' [Eq. (27)] validates the existence of an elasticity principal coordinate system.

It is interesting to note that due to the isotropy of the plane transverse to the fiber direction, every set of axes with one base vector parallel to $\mathbf{x}_3$ is an elasticity principal coordinate system. However, since $\mathbf{x}_1^p$ is chosen along the $\mathbf{x}_1$ direction, a single coordinate system can be identified as R^p.

TABLE II. Reconstruction of both the angular parallax δ and the stiffness tensor (C_{IJ}^p) from the phase velocities measured in the planes $(\mathbf{x}_1',\mathbf{x}_2')$, $(\mathbf{x}_1',\mathbf{x}_3')$, $(\mathbf{x}_1',(45°)')$, and $(\mathbf{x}_1',(135°)')$ for a carbon epoxy composite. As in all the tables, the elasticity constants $(C_{IJ})^{\text{computed}}$ calculated using the stiffness transformation laws are reported in italics. The confidence intervals are associated with the optimal stiffness tensor identified from the experimental data.

Data type	ϕ (degree)	C_{11}^p	C_{12}^p	C_{13}^p	C_{22}^p	C_{23}^p	C_{33}^p	C_{44}^p	C_{55}^p	C_{66}^p
						(GPa)				
Invariant	−29.3	*12.15*	*5.4*	*5.5*	*12.2*	*5.3*	*131*	*6.9*	*5.96*	*3.50*
min $A(\delta)$	−29.6 (0.12)	*12.15*	*5.4*	*5.5*	*12.2*	*5.3*	*131*	*6.9*	*5.96*	*3.50*
min $B((C_{IJ}^p),\delta)$	−29.6 (0.2)	12.15 (0.03)	5.5 (0.1)	5.5 (0.2)	12.3 (0.2)	5.7 (0.7)	131 (3)	6.9 (0.1)	5.97 (0.04)	3.49 (0.04)

IV. APPLICATION TO MATERIALS THAT POSSESS NO KNOWN PLANE OF SYMMETRY

When no assumption about the symmetric character of the observation coordinate system R can be made, the stiffness tensor (C_{IJ}) takes the form (2). However, the material symmetry can be higher than the triclinic symmetry. The successive search for the monoclinic R^m coordinate system and an elasticity principal R^p coordinate system must be carried out.

A. Triclinic media characterization: Numerical simulation

To illustrate the efficiency of the R^p identification, the effect of the initial guesses and random scatter in the experimental data on the results of the functional $B(\mathbf{Y})$'s minimization is considered. The nine moduli (C_{IJ}^p) [Eq. (8)] of the carbon epoxy composite are selected to be the elastic properties to be reconstructed. The unknowns δ^p to be identified are chosen equal to $(-10°, -30°, 0°)$. Let R be the nonprincipal frame obtained from a 30° rotation of R^p about the $\mathbf{x}_1^p$ axis, then through an angle 10° about the transformed $\mathbf{x}_3^p$ axis. R is chosen as the observation coordinate system related to the sample [Fig. 1(a)]. Sets of synthetic velocities are calculated from a set of 21 original elasticity constants (C_{IJ}) relative to R (Table III). The components of (C_{IJ}) are obtained from (C_{IJ}^p) and δ^p using the stiffness transformation laws. The angular ranges experimentally accessible have been accurately determined[1] using the calculation of the transmission coefficient of a signal, propagating once through the sample and taking into account the temporal resolution of the Hilbert transform. Figure 10 shows the phase velocities simulated without scatter in the planes $(\mathbf{x}_1,\mathbf{x}_2)$, $(\mathbf{x}_1,\mathbf{x}_3)$, $(\mathbf{x}_1,45°)$, and $(\mathbf{x}_1,135°)$. The slowness curves are plotted for every propagation direction $\mathbf{n}$ with an associated energy flow vector pointing into the sample.[28,29] For a homogeneous bulk wave propagating in an elastic medium, the energy flow vector is collinear[2] to the normal to the associated slowness surface. The directions $\mathbf{n}$ must satisfy the inequality

$$\frac{\partial f_p}{\partial n_1}(V_p(\mathbf{x}_1),C_{IJ})\,\frac{\partial f_p}{\partial n_1}(V_p(\mathbf{n}),C_{IJ})>0, \qquad (29)$$

where n_1 is the projection of $\mathbf{n}$ upon the normal to the interface $\mathbf{x}_1$. The R^p reconstructions are performed from phase velocity sets simulated with different predetermined scatter levels; 0%, 0.5%, 1%, 1.5%, and 2%.

Up to 1% level, the elasticity constants (C_{IJ}) identified from the simulated data coincide with the original data to the

TABLE III. Results (in GPa) of the elasticity constant reconstruction using simulated velocity data in $\cup$(4 planes) with and without scatter for a carbon epoxy composite studied in a nonprincipal coordinate system.

Data type	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{22}	C_{23}	C_{24}	C_{25}	C_{26}	C_{33}	C_{34}	C_{35}	C_{36}	C_{44}	C_{45}	C_{46}	C_{55}	C_{56}	C_{66}
Original data	12.25	5.65	6.04	0.37	0.44	0.29	21.80	24.58	13.71	2.04	1.44	81.30	36.23	6.39	3.37	25.68	3.58	2.23	6.00	1.43	4.32
Results Scatter 0%	12.25 (0.0)	5.65 (0)	6.04 (0)	0.37 (0)	0.44 (0)	0.29 (0)	21.80 (0)	24.58 (0)	13.71 (0)	2.04 (0)	1.44 (0)	81.30 (0)	36.23 (0)	6.39 (0)	3.37 (0)	25.68 (0)	3.58 (0)	2.23 (0)	6.00 (0)	1.43 (0)	4.32 (0)
Scatter 0.5%	12.25 (0.03)	5.6 (0.1)	6.1 (0.1)	0.37 (0.07)	0.45 (0.02)	0.29 (0.02)	21.6 (0.5)	24.4 (0.6)	13.6 (0.3)	1.99 (0.07)	1.42 (0.07)	81.2 (0.8)	36.2 (0.4)	6.3 (0.2)	3.4 (0.2)	25.7 (0.3)	3.54 (0.06)	2.21 (0.07)	5.98 (0.03)	1.44 (0.03)	4.33 (0.05)
Scatter 1%	12.28 (0.06)	5.7 (0.2)	6.2 (0.2)	0.34 (0.13)	0.47 (0.04)	0.31 (0.04)	21.7 (0.8)	24.2 (1.0)	13.4 (0.6)	2.04 (0.14)	1.40 (0.13)	80.2 (1.5)	35.3 (0.8)	6.4 (0.3)	3.3 (0.3)	25.1 (0.6)	3.54 (0.12)	2.15 (0.14)	5.97 (0.06)	1.49 (0.07)	4.35 (0.10)
Scatter 1.5%	12.34 (0.10)	5.6 (0.3)	6.3 (0.3)	0.39 (0.20)	0.45 (0.06)	0.30 (0.06)	20.7 (1.2)	22.7 (1.5)	12.7 (0.8)	1.87 (0.21)	1.32 (0.20)	78.2 (2.2)	34.3 (1.2)	6.1 (0.4)	3.2 (0.4)	24.5 (0.8)	3.37 (0.17)	2.05 (0.21)	6.00 (0.09)	1.42 (0.11)	4.38 (0.15)
Scatter 2%	12.32 (0.13)	5.4 (0.4)	6.4 (0.4)	0.45 (0.26)	0.46 (0.07)	0.29 (0.08)	19.7 (1.5)	21.7 (2.0)	12.0 (1.0)	1.71 (0.26)	1.18 (0.25)	77.2 (2.8)	33.79 (1.4)	5.7 (0.5)	3.1 (0.5)	24.3 (1.0)	3.23 (0.22)	1.98 (0.26)	5.94 (0.12)	1.45 (0.14)	4.62 (0.20)
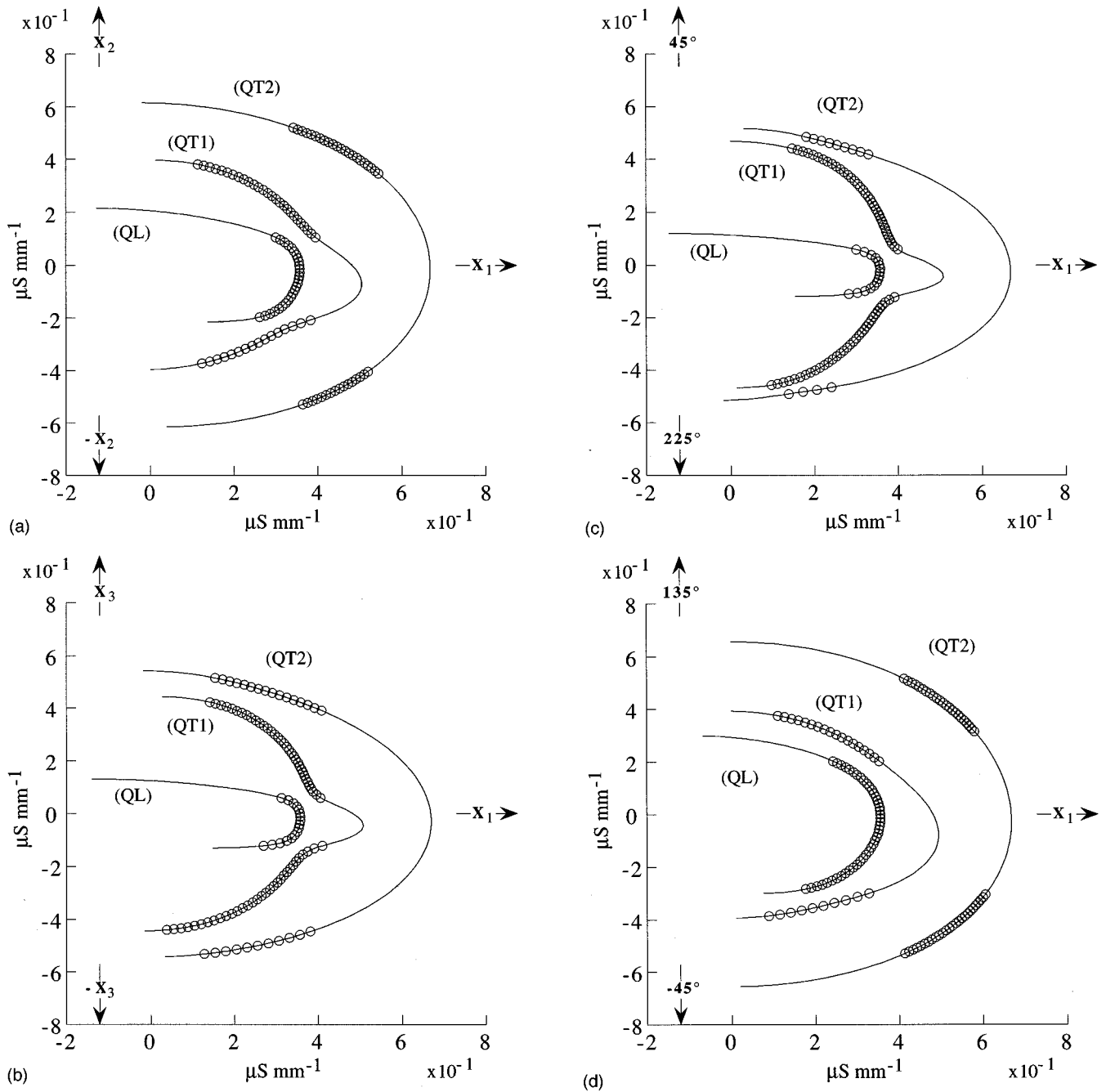
FIG. 10. Slownesses for a medium with apparent triclinic symmetry in the planes (a) $(\mathbf{x}_1, \mathbf{x}_2)$, (b) $(\mathbf{x}_1, \mathbf{x}_3)$, (c) $(\mathbf{x}_1, \mathbf{45°})$, and (d) $(\mathbf{x}_1, \mathbf{135°})$. The points are synthetic slowness data without random scatter. The solid lines are slownesses calculated from the reconstructed elasticity constants.

nearest confidence intervals (Table III). The angular parallax δ^p extracted from the (V_{IJ}) and (d_{IJ}) eigenvectors deduced from (C_{IJ}) are shown in Fig. 11 as a function of the scatter level. This angle determination does not depend continuously on the scatter level. Since the tensors (V_{IJ}) and (d_{IJ}) are built from linear combinations of the stiffnesses (C_{IJ}) without taking into account uncertainties of the elasticity constants, the angular parallax depends strongly both on error propagation in calculations and on the distribution of the identified coefficients (C_{IJ}). This explains the random behavior of this identification as a function of the scatter level. Note that two pairs of closest eigenvectors of (V_{IJ}) and (d_{IJ}) make angles of about 25° for 0.5% scatter and 16° for 1.5%

scatter, whereas the solid angle concerning the last eigenvector couple is very small in the two cases. For 0.5% and 1% levels of scatter in the data, the calculation of the stiffness tensor in the frame located by the angular parallax δ using the stiffness transformation laws is faulty due to the erroneous identification of the angles δ.

The results of the functional $A(\delta)$'s minimization, also reported in Fig. 11, are better. For every scatter level the original and identified angles coincide to the nearest confidence intervals. The general shape of the slowness curves is therefore strongly sensitive to the angular parallax and allows one the accurate location of R^p. The confidence interval set increases as the scatter level in the simulated data in-
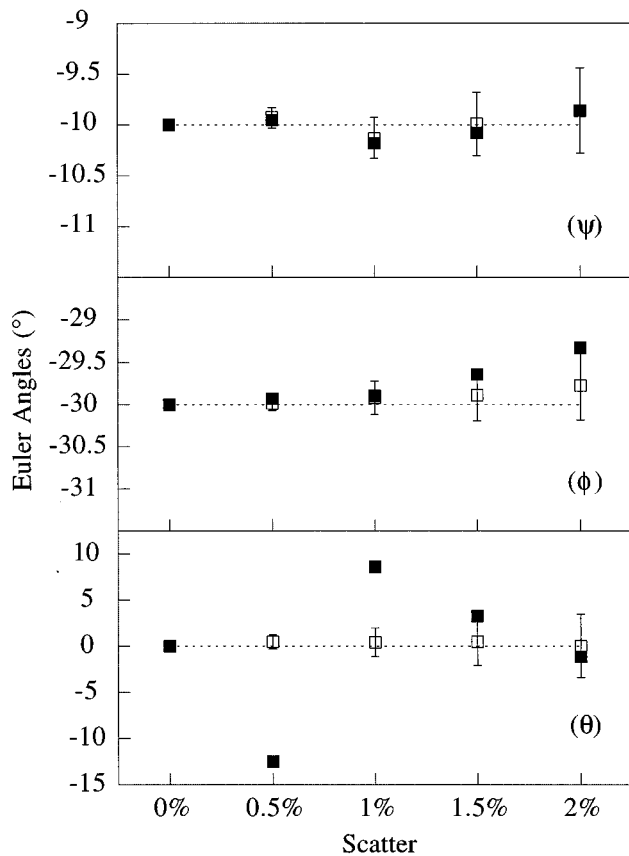
FIG. 11. Stability of the angular parallax (ψ, ϕ, θ) reconstruction against random scatter in the synthetic data simulated in the $(\mathbf{x}_1, \mathbf{x}_2)$, $(\mathbf{x}_1, \mathbf{x}_3)$, $(\mathbf{x}_1, 45°)$, and $(\mathbf{x}_1, 135°)$ planes for a carbon epoxy composite. Determination of ψ, ϕ, and θ using the eigenvectors of the Voigt tensor (V_{IJ}) and of the dilatational modulus (d_{IJ}) (solid squares), and using the functional $A(\delta)$ (empty squares). The original values of angular parallax are shown by dashed lines.

creases. The statistic parameters $I(\delta^p)$ quantify the sensitivity of the algorithm for identifying R^p from wave speed data.

The stability of the $B(\mathbf{Y})$'s minimization against the sets of simulated data and the effect of initial guesses is investigated (Table IV). For each scatter level, the vector δ^p minimizing the functional $A(\delta)$ and the nine moduli $(C_{IJ}^p)^{\mathrm{computed}}$ [deduced from δ^p and (C_{IJ})] are selected to be an initial guess. The other cases of initial guesses are fixed arbitrarily for the angular parallax and determined by adding to each of the nine original moduli (Table IV) an amount defined either arbitrarily or by a percentage ($\pm 30\%$) of its value. The identified solution is unchanged whatever the initial guesses. The identification of R^p is also weakly sensitive to the level of scatter in phase velocity data. Whatever the scatter level, the identified angular parallax δ^p and the original angles coincide to the nearest confidence intervals. The (C_{IJ}^p) reconstruction is more sensitive to the scatter level. The domain of application of this reconstruction method (less than 1%) is identical to the one for the (C_{IJ}) determination (Table III). Since experiments cause a scatter level less than 1%,[1,21,27] satisfactory experimental reconstruction of R^p can be achieved with the proposed procedure. The value of each confidence interval with respect to the scatter level shows that the set of $I(\mathbf{Y})$ quantify *a posteriori* the sensitivity of the inversion algorithm for identifying simultaneously Euler angles and elasticity constants from wavespeed data.

B. Triclinic media characterization: Experimental results

To validate the experimental reconstruction of material symmetry, the experiment is done on the carbon epoxy sample whose fibrous reinforcement direction is known (along the direction $\mathbf{x}_3$). The chosen observation coordinate

TABLE IV. Results of the simultaneous reconstruction of δ and of the stiffness tensor (C_{IJ}^p) using simulated velocity data in $\cup$(4 planes) with and without scatter for a carbon epoxy composite studied in a nonprincipal coordinate system. According to the proposed method, for each scatter level, the results of the functional $A(\delta)$'s minimization; δ (Fig. 11), and the associated stiffness tensor $(C_{IJ})^{\mathrm{computed}}$ are selected to be an initial guess.

Data type	ψ	ϕ	θ	C_{11}^p	C_{12}^p	C_{13}^p	C_{22}^p	C_{23}^p	C_{33}^p	C_{44}^p	C_{55}^p	C_{66}^p
		(degree)						(GPa)				
Original data	-10.0	-30.0	0.0	12.15	5.43	5.46	12.31	5.68	130.28	6.82	6.00	3.48
30%	0	-20	-10	15.80	7.06	7.10	16.00	7.38	169.36	8.87	7.80	4.52
Initial -30%	-20	-40	-10	8.51	3.80	3.82	8.62	3.98	91.20	4.77	4.20	2.44
guesses	-25	-15	15	20	15	15	20	15	160	20	20	20
from min $A(\delta)$												
Results												
Scatter 0%	-10.0	-30.0	0.0	12.15	5.43	5.46	12.31	5.68	130.28	6.82	6.00	3.48
	(0.0)	(0.0)	(0.0)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Scatter 0.5%	-9.9	-30.0	0.6	12.16	5.45	5.5	12.3	5.6	130	6.83	6.00	3.47
	(0.1)	(0.2)	(0.9)	(0.03)	(0.05)	(0.2)	(0.2)	(0.5)	(2)	(0.03)	(0.03)	(0.03)
Scatter 1%	-10.1	-29.9	0.7	12.19	5.52	5.6	12.4	5.7	128	6.86	6.01	3.45
	(0.2)	(0.3)	(1.8)	(0.06)	(0.11)	(0.3)	(0.3)	(0.9)	(3)	(0.06)	(0.06)	(0.07)
Scatter 1.5%	-10.0	-29.9	0.4	12.24	5.49	5.7	12.3	5.5	124	6.82	6.04	3.51
	(0.4)	(0.4)	(2.8)	(0.09)	(0.16)	(0.4)	(0.4)	(1.3)	(4)	(0.09)	(0.09)	(0.11)
Scatter 2%	-9.9	-29.8	-1.6	12.24	5.50	5.7	12.4	5.3	123	6.83	6.10	3.50
	(0.5)	(0.5)	(4.0)	(0.13)	(0.20)	(0.5)	(0.5)	(1.6)	(5)	(0.12)	(0.12)	(0.13)

system $R'' = (\mathbf{x}_1'', \mathbf{x}_2'', \mathbf{x}_3'')$ is obtained from a 30° rotation of R about the $\mathbf{x}_1$ axis followed by a 10° rotation about the transformed $\mathbf{x}_3''$ axis ($\mathbf{x}_3'$) (Fig. 12). The planes $(\mathbf{x}_1'', \mathbf{x}_2'')$, $(\mathbf{x}_1'', \mathbf{x}_3'')$, and $(\mathbf{x}_2'', \mathbf{x}_3'')$ are no longer planes of symmetry.

However, it is difficult to simulate such a frame R'' of propagation in the material from a rotation that is predetermined and applied to the apparatus emitting the ultrasonic waves. The base vector $\mathbf{x}_1''$ is indeed different from the normal to the sample $\mathbf{x}_1$ and according to Snell–Descartes'

laws[2] corresponds to several incident waves. These difficulties disappear as soon as the normal to a sample with triclinic symmetry coincides with a base vector of the observation coordinate system R''.

The study is carried out from the experimental velocities measured for positive and negative incident angles in the data planes $(\mathbf{x}_1, \mathbf{30°})$, $(\mathbf{x}_1, \mathbf{120°})$, $(\mathbf{x}_1, \mathbf{75°})$, and $(\mathbf{x}_1, \mathbf{165°})$ (Fig. 8). The 21 components of the equivalent stiffness tensor in R'',

$$
(C_{IJ}'') = \begin{pmatrix}
12.24 \pm 0.02 & 5.4 \pm 0.2 & 6.2 \pm 0.2 & 0.27 \pm 0.07 & 0.43 \pm 0.02 & 0.28 \pm 0.02 \\
 & 20.9 \pm 0.5 & 23.9 \pm 0.5 & 13.1 \pm 0.3 & 1.9 \pm 0.1 & 1.3 \pm 0.1 \\
 & & 83.4 \pm 1.0 & 36.4 \pm 0.5 & 6.4 \pm 0.2 & 3.2 \pm 0.2 \\
 & & & 25.4 \pm 0.3 & 3.5 \pm 0.1 & 2.2 \pm 0.1 \\
 & \text{Sym.} & & & 5.9 \pm 0.1 & 1.43 \pm 0.03 \\
 & & & & & 4.4 \pm 0.1
\end{pmatrix} \text{(GPa),} \tag{30}
$$

are determined from these data collected in R and then related to R''. The smallest values of the confidence intervals validate *a posteriori* the choice of the data planes. Figure 13 shows the slowness surfaces calculated from the reconstructed moduli for each Christoffel equation's solution in velocity, and the experimental measurements considered for characterization. According to Fig. 9(c), where it can be seen that the slow quasitransverse mode has not been measured in the plane $(\mathbf{x}_1, \mathbf{75°})$, note that only the data of the three other planes are represented in one of the calculated slowness surfaces [Fig. 13(c)]. The calculation of the $T(\theta_i)$ transmission coefficient from the 13 identified elasticity constants (C_{IJ}') [Eq. (27)] shows that the range of the incident angle where the QT2 mode is generated in the plane $(\mathbf{x}_1', (\mathbf{45°})')$ [$= (\mathbf{x}_1, \mathbf{75°})$] is [40°,50°]. This domain is not accessible with our experimental device.

To identify the material symmetry, the normals to the planes of symmetry are independently located. If at least three directions exist and are mutually nearly orthogonal, then they must coincide with a principal coordinate system.

From the components of (C_{IJ}''), the tensors (V_{IJ}) and (d_{IJ}) are determined and their eigenvectors $\mathbf{V}_I$ and $\mathbf{d}_I$ ($I = 1$ to 3) are calculated:

$$
\mathbf{V} = \left\{ \begin{pmatrix} 0.08 \\ 0.48 \\ 0.87 \end{pmatrix}, \begin{pmatrix} 0.15 \\ 0.86 \\ -0.49 \end{pmatrix}, \begin{pmatrix} 0.98 \\ -0.17 \\ 0.00 \end{pmatrix} \right\},
$$

$$
\mathbf{d} = \left\{ \begin{pmatrix} 0.08 \\ 0.48 \\ 0.87 \end{pmatrix}, \begin{pmatrix} 0.15 \\ 0.86 \\ -0.49 \end{pmatrix}, \begin{pmatrix} 0.98 \\ -0.17 \\ 0.00 \end{pmatrix} \right\}. \tag{31}
$$

The angles between the vectors of each couple $(\mathbf{V}_I, \mathbf{d}_I)$ ($I = 1$ to 3) are closest to 0°. Consequently, every average direction $\mathbf{V}_I^a$ determined from the three couples of eigenvectors of closest orientation (Fig. 4) can correspond to the searched normal $\mathbf{x}_1^m$ to a plane of symmetry. The plot of the functional

$A(\phi, \theta)$ (Fig. 14) also emphasizes three minima that can coincide with the normal $\mathbf{x}_1^m$ to be identified. The known material symmetry (quasihexagonal) explains the presence of three probable solutions. First, the monoclinic frame R^m is located with respect to R'' with the two angles (ϕ, θ) (Fig. 5). Since a single direction $\mathbf{x}_1^m$ is useful to search for R^m, only the location of the (known) reinforcement direction is detailed (Table V). The two other solutions are summarized in Table VI. The only modulus C_{11}^m associated with the direction $\mathbf{x}_1^m$ to be determined is reported in these tables.

The unknowns (ϕ, θ) are extracted from the components of the average vector $\mathbf{V}_1^a$ of the eigenvectors pair $(\mathbf{V}_1, \mathbf{d}_1)$:

$$
\mathbf{V}_1^a = 0.08\mathbf{x}_1'' + 0.48\mathbf{x}_2'' + 0.87\mathbf{x}_3''
$$
$$
= \cos(\theta)\mathbf{x}_1'' + \sin(\theta)\cos(\phi)\mathbf{x}_2'' + \sin(\theta)\sin(\phi)\mathbf{x}_3''. \tag{32}
$$

The thirteen components of the equivalent stiffness tensor $(C_{IJ}^m)^{\text{computed}}$ in R^m are calculated from (C_{IJ}'') and the δ^m assessments deduced from the (V_{IJ}) and (d_{IJ}) eigenvectors, then from the minimization of $A(\delta)$. The smaller the angle between the vectors $\mathbf{V}_1$ and $\mathbf{d}_1$ is, the closer to the intrinsic
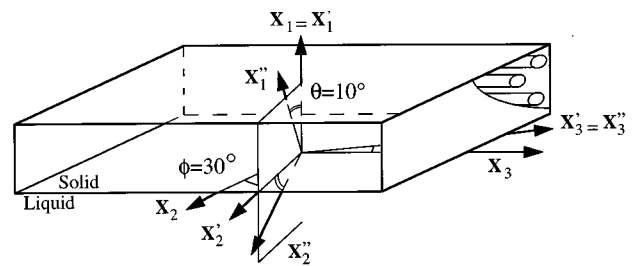
FIG. 12. Choice of a nonprincipal observation coordinate system R'' that does not coincide with the geometric coordinate system R of a unidirectional carbon epoxy composite. R'' is obtained from R by two rotations $\phi = 30°$ and $\theta = 10°$.

(a)

(b)

(c)
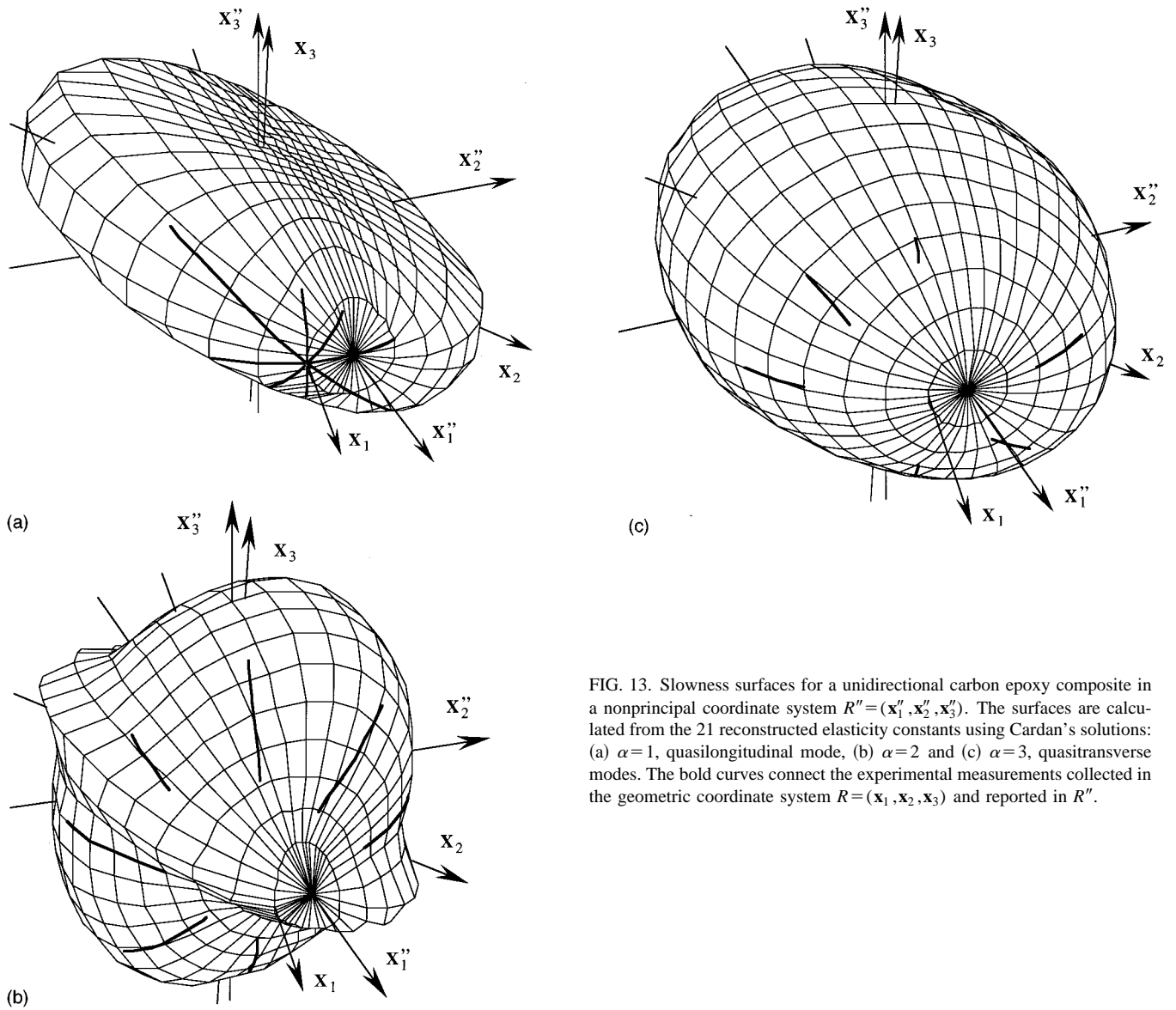
FIG. 13. Slowness surfaces for a unidirectional carbon epoxy composite in a nonprincipal coordinate system $R'' = (\mathbf{x}_1'', \mathbf{x}_2'', \mathbf{x}_3'')$. The surfaces are calculated from the 21 reconstructed elasticity constants using Cardan's solutions: (a) $\alpha = 1$, quasilongitudinal mode, (b) $\alpha = 2$ and (c) $\alpha = 3$, quasitransverse modes. The bold curves connect the experimental measurements collected in the geometric coordinate system $R = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ and reported in R''.

solution the estimate deduced from the average direction $\mathbf{V}_1^a$ is. This explains that the vector δ^m determined from the $\mathbf{V}_1^a$ direction and the associated tensor $(C_{IJ}^m)^{\text{computed}}$ form a fairly good estimate of the optimal result. To validate the existence of R^m, the stiffness tensor $(C_{IJ}'')^{\text{computed}}$ equivalent to the identified solution $((C_{IJ}^m), \delta^m)$ in R'' is calculated:

$$(C_{IJ}'')^{\text{computed}} =$$

$$\begin{vmatrix} 12.25 & 5.65 & 6.04 & 0.36 & 0.43 & 0.29 \\ & 21.46 & 24.26 & 13.40 & 1.99 & 1.39 \\ & & 82.61 & 36.28 & 6.40 & 3.32 \\ & & & 25.43 & 3.54 & 2.18 \\ \text{Sym.} & & & & 5.99 & 1.40 \\ & & & & & 4.29 \end{vmatrix} \text{(GPa)}.$$

(33)

The components of this tensor and the 21 elasticity constants identified from the experimental data [Eq. (30)] coincide, to

the nearest confidence intervals. The slight deviation Δ ($=1.1\%$) between these two tensors confirms the existence of the monoclinic frame R^m. The unit vector $\mathbf{x}_1^m$ defined by $(\phi^m, \theta^m) = (60.8°, 85.1°)$ is therefore a normal to a plane of symmetry. The vector $\mathbf{x}_1^m$ is expressed in R and R^p (reconstructed in Sec. III B) by, respectively,

$$\mathbf{x}_1^m = 0.00\mathbf{x}_1 - 0.01\mathbf{x}_2 + 1.00\mathbf{x}_3,$$

$$\mathbf{x}_1^m = 0.00\mathbf{x}_1^p + 0.00\mathbf{x}_2^p + 1.00\mathbf{x}_3^p.$$

(34)

This vector is aligned with the base vector $\mathbf{x}_3^p$ of the frame R^p determined in Sec. III B. This is confirmed by the comparison of the elasticity constants $C_{33}^p = 131$ GPa (Table II) and $C_{11}^m = 131$ GPa (Table V). The experimental location of a normal to a plane of symmetry is accurately achieved.

Two other normals to a plane of symmetry were able to be identified from the two average vectors $\mathbf{V}_2^a$ and $\mathbf{V}_3^a$ (Table VI). They are respectively collinear to the base vectors $\mathbf{x}_2^p$ and $\mathbf{x}_1^p$ of the elasticity principal frame R^p identified in Sec. III B. The agreement between the moduli C_{11}^m associated
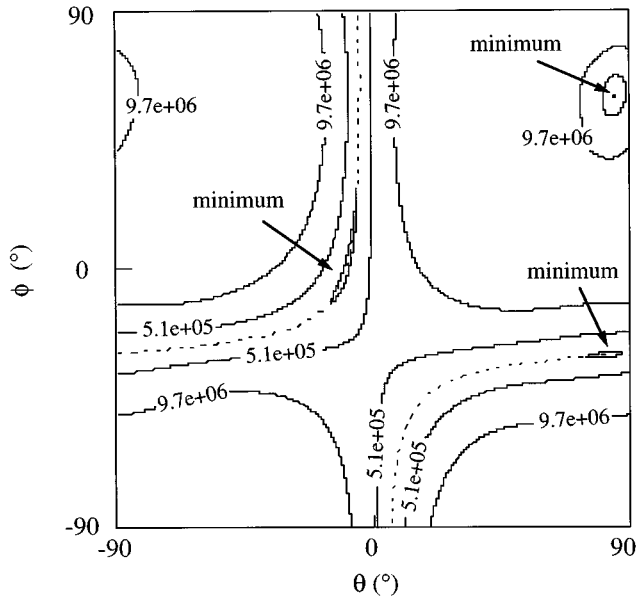
FIG. 14. Isovalues of the functional $A(\delta)$ calculated from the equivalent stiffness tensor (C''_{IJ}) in the observation frame R'' and the phase velocities measured in the $(\mathbf{x}'_1,\mathbf{x}'_2)$, $(\mathbf{x}'_1,\mathbf{x}'_3)$, $(\mathbf{x}'_1,(45°)')$, and $(\mathbf{x}'_1,(135°)')$ planes for a carbon epoxy composite. Three minima are clearly visible. The dashed lines are formed by the directions lying in the reconstructed isotropic plane.

TABLE VI. Reconstruction of three normals to a plane of symmetry from the phase velocities measured in the planes $(\mathbf{x}'_1,\mathbf{x}'_2)$, $(\mathbf{x}'_1,\mathbf{x}'_3)$, $(\mathbf{x}'_1,(45°)')$, and $(\mathbf{x}'_1,(135°)')$ for a carbon epoxy composite.

Data type	ϕ (degree)	θ (degree)	C_{11}^m (GPa)
From $\mathbf{V}_1^a$	60.8 (0.1)	85.1 (0.1)	131 (2)
From $\mathbf{V}_2^a$	−30.0 (0.2)	81.3 (0.7)	12.3 (0.2)
From $\mathbf{V}_3^a$	0.0 (1.5)	−10.0 (0.5)	12.15 (0.02)

with these identified directions and the elasticity constants $C_{22}^p(=12.3$ GPa$)$ and $C_{11}^p(=12.15$ GPa$)$ (Table II) is good. Therfore, three normals to the planes of symmetry have been reconstructed independently from each other.

Let us search for the elasticity principal coordinate system R^p located with respect to the observation frame R'' by the three Euler angles $\delta=(\psi,\phi,\theta)$. The three average directions $\mathbf{V}_1^a$, $\mathbf{V}_2^a$, and $\mathbf{V}_3^a$ must superimpose on the crystallographic axes. A reference triad can be located by several combinations of Euler angles. To identify the closest angles triplet δ^p to the "simulated" angles $(-10°,-30°,0°)$, the basis $(\mathbf{V}_2^a,-\mathbf{V}_3^a,\mathbf{V}_1^a)$ is chosen to be the first assessment of R^p. The results of the successive stages of the R^p reconstruction are summarized in Table VII. Using the rotation matrix $(P_{IJ})_\delta$ associated with the transformation between the triads $(\mathbf{x}''_1,\mathbf{x}''_2,\mathbf{x}''_3)$ and $(\mathbf{V}_2^a,-\mathbf{V}_3^a,\mathbf{V}_1^a)$, a first estimate of δ^p is obtained. This solution $\delta^p[=(-10.0,-29.2,0.0)]$ and the associated stiffness tensor $(C_{IJ}^p)^{\text{computed}}$ are very close to the optimal solution. This is explained by the accurate reconstruction of the elasticity tensor (C''_{IJ}) [Eq. (30)]. However,

TABLE V. Reconstruction of the angular parallax δ and the modulus C_{11}^m from the phase velocities measured in the planes $(\mathbf{x}'_1,\mathbf{x}'_2)$, $(\mathbf{x}'_1,\mathbf{x}'_3)$, $(\mathbf{x}'_1,(45°)')$, and $(\mathbf{x}'_1,(135°)')$ for a carbon epoxy composite.

Data type	ϕ (degree)	θ (degree)	C_{11}^m (GPa)
Eigenvectors	61.2	85.1	*131*
min $A(\delta)$	60.8 (0.1)	85.1 (0.1)	*131*
min $B((C_{11}^m),\delta)$	60.8 (0.1)	85.1 (0.1)	131 (2)

this study has been carried out assuming the existence of R^p. To verify this assumption, the stiffness tensor $(C''_{IJ})^{\text{computed}}$ equivalent to the optimal solution $((C_{IJ}^p),\delta^p)$ in R'' is calculated. The components of $(C''_{IJ})^{\text{computed}}$ are identical to the ones of the tensor (33). Therefore the deviation Δ between these moduli and (C''_{IJ}) is unchanged (1.1%). The frame located with $(\psi^p,\phi^p,\theta^p)=(-10.0°,-29.6°,0.0°)$ is the elasticity principal coordinate system R^p. The base vectors of R^p are expressed in the geometric frame R by

$$\mathbf{x}_1^p=1.00\mathbf{x}_1+0.00\mathbf{x}_2+0.00\mathbf{x}_3,$$
$$\mathbf{x}_2^p=0.00\mathbf{x}_1+1.00\mathbf{x}_2+0.01\mathbf{x}_3, \qquad (35)$$
$$\mathbf{x}_3^p=0.00\mathbf{x}_1-0.01\mathbf{x}_2+1.00\mathbf{x}_3.$$

This orthonormal triad $(\mathbf{x}_1^p,\mathbf{x}_2^p,\mathbf{x}_3^p)$ superimposes on the elasticity principal coordinate system indentified in Sec. III B. The optimal tensor (C_{IJ}^p) (determined simultaneously with δ^p, Table V) is identical to the one already established in Sec. III B, Table II. The coincidence between the results of these two identifications is due to the taking into account of the same set of experimental data. The experimental identification of the principal coordinate system R^p [Eq. (35)] can therefore be carried out from only the wave speed measurements.

However, the material symmetry can be higher than the orthorhombic symmetry. The classical symmetries occurring in composite materials are successively considered. For example, the assumption of hexagonal symmetry is described by the following equivalent tensor in R^p [Eq. (26)]:

$$(C_{IJ}^{\text{Hex}})=$$

$$\begin{pmatrix} 12.27 & 5.40 & 5.57 & & & \\ & 12.27 & 5.57 & & 0 & \\ & & 130.61 & & & \\ & & & 6.43 & & \\ \text{Sym.} & & & & 6.43 & \\ & & & & & 3.43 \end{pmatrix} \text{(GPa)}.$$

$$(36)$$

The deviations from isotropic, hexagonal, and tetragonal models [Eq. (25)] are shown in Fig. 15 with the ones already calculated at the time of the search for the frames $R^m(\Delta=1.1\%)$ and $R^p(\Delta=1.1\%)$. Clearly it appears that the more

TABLE VII. Reconstruction of both the angular parallax δ and the stiffness tensor (C_{IJ}^p) from the phase velocities measured in the planes ($\mathbf{x}_1'$,$\mathbf{x}_2'$), ($\mathbf{x}_1'$,$\mathbf{x}_3'$), ($\mathbf{x}_1'$, $(45°)'$) and ($\mathbf{x}_1'$, $(135°)'$) for a carbon epoxy composite.

Data type	ψ	ϕ	θ	C_{11}^p	C_{12}^p	C_{13}^p	C_{22}^p	C_{23}^p	C_{33}^p	C_{44}^p	C_{55}^p	C_{66}^p
	(degree)			(GPa)								
Eigenvectors	−10.0	−29.2	0.0	12.15	5.4	5.5	12.2	5.3	131	6.9	5.96	3.50
min $A(\delta)$	−10.0 (0.1)	−29.6 (0.1)	0.0 (0.6)	12.15	5.4	5.5	12.2	5.3	131	6.9	5.96	3.50
min $B((C_{IJ}^p),\delta)$	−10.0 (0.1)	−29.6 (0.1)	0.0 (0.6)	12.15 (0.02)	5.5 (0.1)	5.5 (0.1)	12.3 (0.2)	5.7 (0.5)	131 (2)	6.9 (0.1)	5.97 (0.03)	3.48 (0.03)

probable material symmetry is the hexagonal symmetry. The classic symmetry of unidirectional composites is well identified. In the plane ($\mathbf{x}_1^p$,$\mathbf{x}_2^p$) the value of the anisotropy factor $I = 2C_{66}^p/(C_{11}^p - C_{12}^p)$ ($=1.04$) or $I = 2C_{66}^p/(C_{22}^p - C_{12}^p)$ ($=1.02$) is very close to the one of the isotropy ($I=1$). The deviation from hexagonal symmetry is therefore slight. Every direction lying in the isotropic plane ($\mathbf{x}_1^p$,$\mathbf{x}_2^p$) must correspond to a normal to a plane of symmetry. Each of these directions can be expressed in the observation coordinate system R'' with the two angles (ϕ,θ) defined in Fig. 5. This set of angles couples is shown in Fig. 14 by dashed lines. If the material symmetry is exactly hexagonal, these two curves as a dashed line must superimpose on the set of minima of the functional $A(\delta)$. Although a decrease of $A(\delta)$ appears around these curves, only two points of these lines minimize the functional $A(\delta)$. The experimental errors on velocity measurements or the quasihexagonal symmetry of the sample can explain the identification of a single principal coordinate system among every frame which has a base vector along the direction $\mathbf{x}_3^p$.

During this experimental characterization, the quality of the wave speed measurements and the accuracy of the simultaneous identification of the angular parallax locating the higher symmetry frame and the associated optimal stiffness tensor are very good. Due to the small solid angle (close to 0°) around the average directions between the three closest eigenvectors of (V_{IJ}) and (d_{IJ}), the initial assessment from these two tensors is excellent. However, a weak level of scatter in the ultrasonic data induces an erroneous identification of the higher symmetry frame using the tensors (V_{IJ})

and (d_{IJ}): the proximity of theirs closest eigenvectors is removed. Others experimental characterizations[30] have demonstrated both the applicability and the advantage of the proposed method.

V. CONCLUSION

A method for identification of material symmetry from ultrasonic measurements has been introduced. By minimizing a functional, both the higher symmetry coordinate system and the elasticity constants are extracted from the experimental velocities of ultrasonic waves obliquely incident from water onto a plate. Numerical as well as experimental results show the effectiveness of the method. The stability of the inversion algorithm has been demonstrated by simulation on data with random scatter. The experimental reliability has been established by means of material symmetry reconstruction from wave speed measurements in an arbitrarily oriented but known composite material.

The location of the higher symmetry coordinate system is weakly sensitive to the level of scatter in phase velocity data. On the other hand, the associated optimal stiffness tensor reconstruction is more sensitive to the scatter level. A typical experimental scatter level less than 1% is suitable. The estimate of confidence intervals quantifies *a posteriori* the sensitivity of the inversion algorithm for identifying simultaneously Euler angles and elasticity constants from wave speed data. A better resolve is obtained by using directly the wave speed measurements, particularly when the complete elasticity tensor is not accurately identified in the observation coordinate system. The limit of the methods that do not explicitly account for the experimental data has been emphasized using simulated data with scatter. When the scatter level in the wave speed measurements is different from 0%, which is naturally the case for an experimental characterization, the only proposed method allows one to identify the higher symmetry frame and the associated stiffness tensor. A confidence interval is then associated both with each identified Euler angle and with each identified elasticity constant. It is as important to identify a physical property as to give an uncertainty of this data.

A weak angular parallax between the observation frame and an elasticity principal coordinate system (due to a possible defective sample positioning) leads to non-negligible errors in the determination of some elasticity constants. This mainly has repercussions on the Poisson ratios. That may
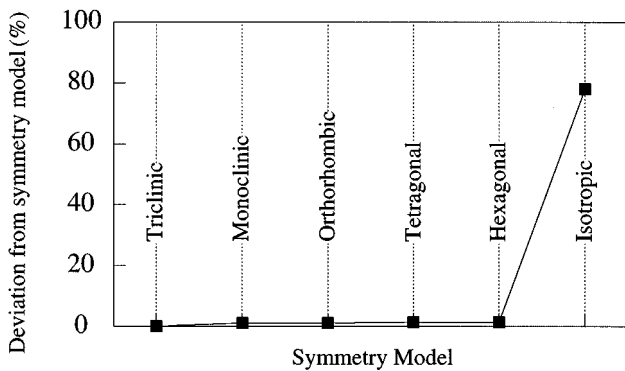
FIG. 15. The more probable symmetry for a unidirectional carbon epoxy composite whose characterization has been performed in a nonprincipal coordinate system.

explain some discordance between the ultrasonic characterization and the mechanical measurements. The reconstruction of the stiffness tensor must be carried out by making the fewest assumptions possible on the coincidence between the observation frame and the symmetry elements. The location of the highest symmetry coordinate system improves the search for the more probable material symmetry.

The present approach can allow one to characterize the wood whose annular rings induce a monoclinic symmetry model or the industrial composite materials whose strata's stacking defects remove or conceal the material symmetry coordinate system. Another attractive feature is that the elastic properties of composite under off-axis loading can also be followed. The problem of damage-induced anisotropy, i.e., the orientation of matrix microcracks comparative to the loading direction and to the natural axes of the composite, may be treated without the limiting assumption that the material keeps its orthorhombic symmetry in an *a priori* known coordinate system.

[1] C. Aristégui and S. Baste, ''Optimal recovery of the elasticity tensor of general anisotropic materials from ultrasonic velocity data,'' J. Acoust. Soc. Am. **101**, 813–833 (1997).

[2] B. A. Auld, *Acoustic Fields and Waves in Solids* (Wiley–Interscience, New York, 1973).

[3] C. M. Sayers, ''Inversion of ultrasonic wave velocity measurements to obtain the microcrack orientation distribution function in rocks,'' Ultrasonics **26**, 73–77 (1988).

[4] M. L. Dunn and H. Ledbetter, ''Estimation of the orientation distribution of short-fiber composites using ultrasonic velocities,'' J. Acoust. Soc. Am. **99**, 283–291 (1996).

[5] W. C. Van Buskirk, S. C. Cowin, and R. Carter, Jr., ''A theory of acoustic measurement of the elastic constants of a general anisotropic solid,'' J. Mater. Sci. **21**, 2749–2762 (1986).

[6] S. C. Cowin and M. M. Mehrabadi, ''On the identification of material symmetry for anisotropic elastic materials,'' Q. J. Mech. Appl. Math. **40**, 451–476 (1987).

[7] S. C. Cowin, ''Properties of the anisotropic elasticity tensor,'' Q. J. Mech. Appl. Math. **42**, 249–266 (1989).

[8] A. N. Norris, ''On the acoustic determination of the elastic moduli of anisotropic solids and acoustic conditions for the existence of symmetry planes,'' Q. J. Mech. Appl. Math. **42**, 412–426 (1989).

[9] J. P. Jaric, ''On the conditions for the existence of a plane of symmetry for anisotropic elastic material,'' Mech. Res. Commun. **21**, 153–174 (1994).

[10] F. E. Borgnis, ''Specific directions of longitudinal waves propagation in anisotropic media,'' Phys. Rev. **98**, 1000–1005 (1955).

[11] I. I. Kolodner, ''Existence of longitudinal waves in anisotropic media,'' J. Acoust. Soc. Am. **40**, 730–731 (1966).

[12] R. Arts, ''A study of general anisotropic elasticity in rocks by wave propagation. Theoretical and experimental aspects,'' Ph.D. thesis, Université Paris VI, 1993.

[13] A. G. Every, ''General closed-form expressions for acoustic waves in elastically anisotropic solids,'' Phys. Rev. B **22**, 1746–1760 (1980).

[14] B. Castagnède, J. T. Jenkins, W. Sachse, and S. Baste, ''Optimal determination of the elastic constants of composite materials from ultrasonic wavespeed measurements,'' J. Appl. Phys. **67**, 2753–2761 (1990).

[15] S. I. Rokhlin and W. Wang, ''Ultrasonic evaluation of in-plane and out-of-plane elastic properties of composite materials,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1989), Vol. 8B, pp. 1489–1496.

[16] S. I. Rokhlin and W. Wang, ''Double through-transmission bulk wave method for ultrasonic phase velocity measurement and determination of elastic constants of composite materials,'' J. Acoust. Soc. Am. **91**, 3303–3312 (1992).

[17] B. Audoin and J. Roux, ''Traitement de signal ultrasonore adapté à l'évaluation non destructive de matériaux dispersifs,'' Rev. Phys. Appl. **25**, 1011–1017 (1990).

[18] B. Audoin and J. Roux, ''An innovative application of the Hilbert transform to time delay estimation of overlapped ultrasonic echoes,'' Ultrasonics **34**, 25–33 (1996).

[19] B. Audoin, ''Evaluation ultrasonore de l'endommagement anisotrope d'un composite à matrice fragile,'' Ph.D. thesis, 620, Université Bordeaux I, 1991.

[20] B. Audoin and S. Baste, ''Ultrasonic evaluation of stiffness tensor changes and associated anisotropic damage in a ceramic matrix composite,'' J. Appl. Mech. **61**, 309–316 (1994).

[21] Y. C. Chu and S. I. Rokhlin, ''Comparative analysis of through-transmission ultrasonic bulk wave methods for phase velocity measurements in anisotropic materials,'' J. Acoust. Soc. Am. **95**, 3204–3212 (1994).

[22] S. Baste and B. Hosten, ''Identification complète de la matrice de raideur par propagation hors plan principal,'' Rev. Phys. Appl. **25**, 161–168 (1990).

[23] T. C. T. Ting, ''Invariants of anisotropic elastic constants,'' Q. J. Mech. Appl. Math. **40**, 431–448 (1987).

[24] A. Gourdin and M. Boumahrat, *Méthodes numériques appliquées* (Lavoisier, Paris, 1983).

[25] M. Minoux, *Programmation mathématique Théorie et algorithms* (Dunod, Paris, 1983).

[26] B. Hosten, ''Elastic characterization of orthotropic composite materials from ultrasonic inspection through non-principal planes,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1991), Vol. 10B, pp. 1437–1444.

[27] Y. C. Chu and S. I. Rokhlin, ''Stability of determination of composite moduli from velocity data in planes of symmetry for weak and strong anisotropies,'' J. Acoust. Soc. Am. **95**, 213–225 (1994).

[28] E. G. Henneke II, ''Reflexion–refraction of a stress wave at a plane boundary between anisotropic media,'' J. Acoust. Soc. Am. **51**, 210–217 (1972).

[29] S. I. Rokhlin, T. K. Bolland, and L. Adler, ''Reflection and refraction of elastic waves on a plane interface between two generally anisotropic media,'' J. Acoust. Soc. Am. **79**, 906–918 (1986).

[30] C. Aristégui, ''Anisotropie élastique totale naturelle ou induite,'' Ph.D. thesis, Université Bordeaux I, 1997.

Bjerknes force and bubble levitation under single-bubble sonoluminescence conditions

Thomas J. Matula, Sean M. Cordry,[a] Ronald A. Roy,[b] and Lawrence A. Crum

Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

Bubble levitation in an acoustic standing wave is re-examined here for the case of single-bubble sonoluminescence. The equilibrium position of the bubble is calculated by equating the average Bjerknes force with the average buoyancy force. The predicted values, as a function of pressure amplitude, are compared with experimental measurements. Our measurements indicate that the equilibrium position of the bubble shifts *away* from the pressure antinode as the drive pressure increases, in qualitative agreement with calculations, but unexpected when only linear theory is considered [A. Eller, J. Acoust. Soc. Am. **43**, 170 (1968)]. The Bjerknes force also provides an upper limit to the drive pressure in which a bubble can be levitated near (and above) a pressure antinode, even in the absence of shape instabilities. © *1997 Acoustical Society of America.* [S0001-4966(97)05209-0]

PACS numbers: 43.35.Ei [HEB]

INTRODUCTION

In 1989, Gaitan discovered that a single air bubble trapped at a pressure antinode in a standing wave could be made to emit light.[1] This discovery of single-bubble sonoluminescence, or SBSL, has led to many remarkable experimental measurements of the light (and sound) emitted from single bubbles.[2–8] Of fundamental importance for generating cavitation bubbles that can remain stable for hours at a time is the process of bubble levitation.

Previous studies of bubble levitation in a stationary wave have been carried out for bubbles larger than resonance size (where the bubble is levitated near a pressure node)[9,10] and for bubbles smaller than resonance size,[11,12] in both cases driven into small amplitude oscillations. In this paper we will consider in more detail the levitation of bubbles used for SBSL experiments (bubbles below resonance size, driven into highly nonlinear oscillations). In particular we will look at the role of the Bjerknes force on a bubble (described later), since it is this force that causes a bubble to be levitated against the gravitational (buoyancy) force. Although little is mentioned about the light emission from a sonoluminescing bubble in this paper, our calculations and measurements were carried out in the parameter space where the stable sonoluminescence of single, air-filled bubbles occur.

Our calculations show that for low driving pressures, a small increase in pressure leads to a larger (in magnitude) Bjerknes force, as expected. This causes the bubble to be pulled closer to the antinode. However, at larger driving pressures, the magnitude of the Bjerknes force actually becomes *smaller*, and the bubble's equilibrium levitation position shifts *away* from the antinode. We will show how this occurs, with experimental measurements of the levitation position of a sonoluminescence bubble that shows qualitative agreement with our calculations. Quantitative agreement can

be obtained if one assumes that the presence of the bubble perturbs the pressure profile in the near vicinity of the bubble. If true, this may have implications for building stable SBSL levitation cells.

In the following sections we make extensive use of the radius-time profile calculated for SBSL bubbles. The bubble's volume mode pulsations are calculated using the Keller equation, given by[13]

$$\left(1 - \frac{\dot{R}}{c}\right) R\ddot{R} + \frac{3}{2}\dot{R}^2\left(1 - \frac{\dot{R}}{3c}\right) = \frac{1}{\rho}\left(1 + \frac{\dot{R}}{c} + \frac{R}{c}\frac{d}{dt}\right)$$
$$\times [p_b - p_a(t)], \quad (1)$$

where

$$p_b = \left(p_0 + \frac{2\sigma}{R_0} - p_v\right)\left(\frac{R_0^{3\kappa}}{(R^3 - a^3)^\kappa}\right) - \frac{2\sigma}{R} - 4\mu\frac{\dot{R}}{R}. \quad (2)$$

We choose parameters that match our laboratory conditions; a viscosity of $\mu = 1$ cP, a surface tension coefficient of $\sigma = 72$ dyn/cm, the sound speed in water $c = 1490$ m/s, a vapor pressure at room temperature of $p_v = 2660$ Pa, and an ambient pressure of $p_0 = 1.013 \times 10^5$ Pa. We also assume that the gas (with $\kappa = 1.4$) has a van der Waals hard core, with $R_0/a = 8.54$, where a is the excluded radius and $R_0 = 5$ μm is the equilibrium radius of the bubble.

We begin by examining the Bjerknes force on a small bubble, which is levitated just above the antinode of a standing wave. Then the equilibrium position of the bubble, assuming the average buoyancy force just balances the Bjerknes force, is considered. Again, we choose parameters that match laboratory conditions for our experimental system, in this case a rectangular levitation cell with dimensions 8.5×8.5-cm cross section, and a water column ($\rho = 1000$ kg/cm³) of 14.6 cm. The drive frequency for this system is 19.5 kHz, being operated in the (1,1,3) mode. That is, there are three pressure antinodes along the vertical direction. We choose one of the antinodes as the origin of our coordinate system. For the bubble to be in stable equilibrium,
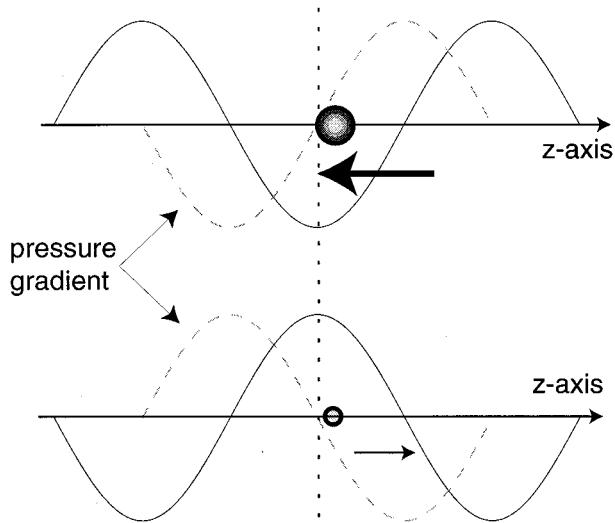
FIG. 1. For small drive pressures, the pressure force attracting the bubble towards the pressure antinode during the first half of the acoustic cycle is greater than the force pushing the bubble away during the second half of the acoustic cycle, since the volume of the bubble is greater in the former case (illustrated by the size of the arrows). Thus over an acoustic cycle, the average force is directed toward the pressure antinode. The solid lines refer to the drive pressure, the dashed lines refer to the pressure gradient, and the vertical line marks the location of the antinode. The z axis corresponds to the vertical axis.

the horizontal radiation forces must be zero, thus we consider only forces in the vertical direction in which gravity acts. Note that with three antinodes, the vertical wavelength is $\lambda = 2/3 \times 14.6 \approx 9.73$ cm.

I. THE BJERKNES FORCE

In an acoustic standing wave, bubbles can be levitated against the gravitational force of buoyancy by the well-known Bjerknes force, $F_{\text{Bjerk}} = \langle F_p \rangle = \langle -V(t)\nabla P \rangle$, where F_p is the instantaneous pressure force on the bubble, ∇P is the gradient of the applied acoustic pressure, and $\langle \cdots \rangle$ denotes the time average.[14–16] In our levitation cell the standing wave pressure profile (along the vertical axis) is given by $P(z,t) = -P_a \cos(k_z z)\sin(\omega t)$. The magnitude of this pressure gradient is then $\nabla P = k_z P_a \sin(k_z z)\sin(\omega t)$, so that

$$F_p = (-4\pi/3)R^3 k_z P_a \sin(k_z z)\sin(\omega t). \tag{3}$$

Here $k_z = 2\pi/\lambda_z$ is the vertical wave number, $\lambda_z = 9.73$ cm is the vertical wavelength, z is the position above the antinode $\omega = 2\pi f$, where f is the drive frequency and P_a is the drive pressure amplitude. Note that the sign of the pressure force changes as the drive pressure changes.

Physically, the Bjerknes force arises from a pressure difference (gradient) across the bubble. Figure 1 illustrates this for the case of small driving pressures (and for drive frequencies below the bubble's natural resonance frequency). During the negative portion of the sound field, the bubble grows. There is a pressure force on the bubble due to a slight difference in pressure exerted on either side of the bubble's surface. This force directs the bubble towards the pressure antinode. During the compressive phase of the sound field, the bubble is small, and the force is directed away from the
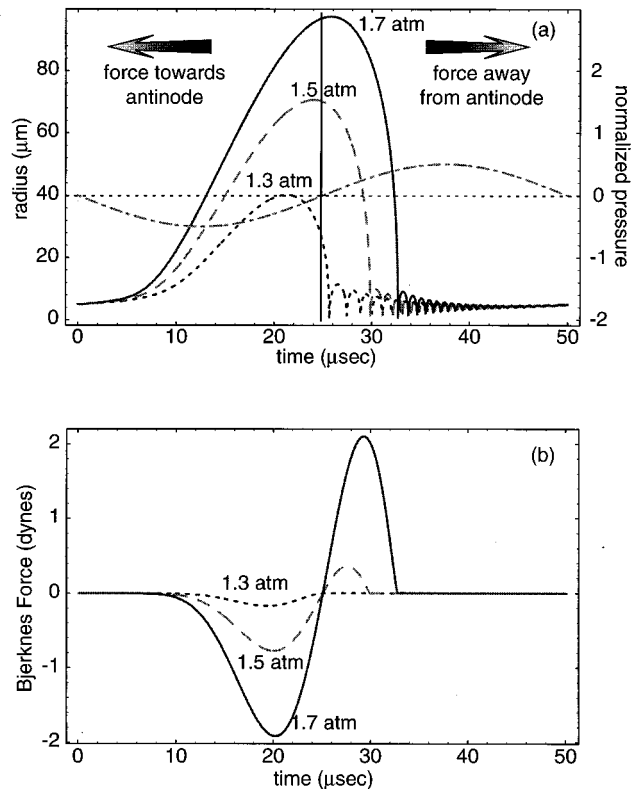
FIG. 2. (a) Calculated steady-state radius-time curves for a 5-μm bubble driven with a pressure amplitude of 1.3, 1.5, and 1.7 atm using Eq. (1). Note that as the pressure amplitude is increased, the bubble collapses later in the acoustic cycle, such that it remains relatively large even after the pressure changes phase. The arrows illustrate the direction of the Bjerknes force, towards the pressure antinode during the first half-cycle, and away from the pressure antinode during the second half cycle. (b) This instantaneous pressure force for bubbles driven as shown in (a).

pressure antinode. However, since the corresponding volume is smaller, this force is smaller, and hence, over an acoustic cycle the average (or Bjerknes) force directs the bubble towards the antinode. This argument on the direction of the force applies only to a bubble that is driven below its natural resonance frequency. For bubbles driven above their natural resonance frequency, a different phase response requires them to be forced away from the pressure antinode and toward a node.

Though valid at lower drive pressure amplitudes, the results of this description of the Bjerknes force must be modified at higher driving pressures. Consider Fig. 2(a) which shows steady-state radius-time curves for a bubble ($R_0 = 5$ μm) with three different driving pressures. The normalized driving pressure is also shown. Recall that the pressure force changes sign with the drive pressure. This demarcation line is shown near $t \approx 25$ μs for the particular parameters used in the calculation. The pressure force is directed towards the pressure antinode during the tensile portion of the sound field, and away from the antinode during the positive portion.

For small drive pressures the volume of the bubble at its maximum size (where it contributes most to the pressure force) occurs during the negative portion of the sound field. However, as the drive pressure increases, the bubble col-
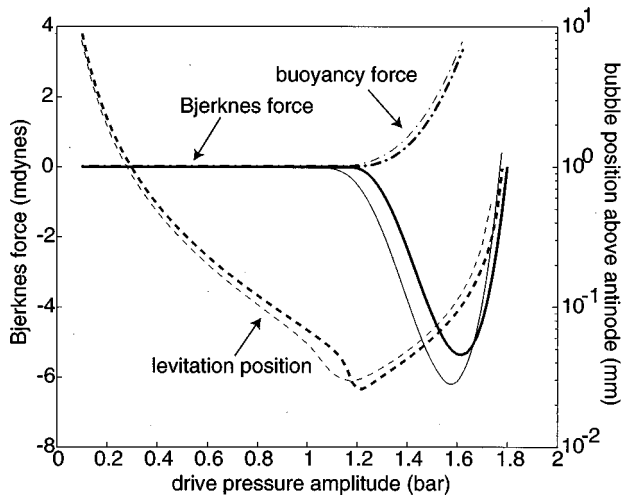
FIG. 3. The Bjerknes force, buoyancy force, and equilibrium bubble position for bubbles with an equilibrium radius of 3 μm (heavy lines) and 5 μm (light lines). For small drive pressures, the Bjerknes force is small and negative. At about 1.78 atm the force becomes positive, and the system can no longer support a stable bubble near the pressure antinode. Note that the equilibrium position of SBSL bubbles (bubbles driven above the lower threshold of 1.2 bar) shifts away from the pressure antinode as the drive pressure is increased. Also note that since the bubble is not far from the antinode, the drive pressure at the bubble is nearly identical to the pressure at the antinode.

lapses later in the cycle and can still be large after the pressure turns positive. With the bubble's volume near its maximum after the pressure turns positive, the pressure force contributes a relatively large component directed away from the pressure antinode. This is illustrated in Fig. 2(b) which shows the instantaneous pressure force calculated for the three R-t curves shown in Fig. 2(a).

The Bjerknes force is the average pressure force, averaged over an acoustic cycle,

$$F_{\mathrm{Bjerk}} = \langle F_p \rangle = \frac{-4\pi}{3T} k_z P_a \sin(k_z z) \int_0^T R(t)^3 \sin(\omega t) dt, \quad (4)$$

where $T = 1/f$ is the acoustic period. This force is plotted in Fig. 3 as a function of drive pressure amplitude (assuming the equilibrium position $z = 1$ mm). For small pressure amplitudes, the Bjerknes force is small and negative (but not zero), directing the bubble towards the pressure antinode. As the pressure amplitude increases beyond 1.2 bar (near the cavitation threshold) the magnitude of the Bjerknes force increases rapidly, resulting from the rapidly increasing volume-mode pulsations of the bubble; however, note that at even higher pressure amplitudes, the slope of the force changes sign, and the magnitude of the Bjerknes force begins to decrease.

The change in slope near 1.6 bar results from the bubble collapsing later in the acoustic cycle, as shown in Fig. 2. The integral in Eq. (4) begins to contribute a large positive component to the average force, which competes with, and eventually dominates the negative components in front of the integral. Eventually (in this case, for $P_a \approx 1.78$ bar) the Bjerknes force becomes positive, and consequently the bubble is pushed away from the antinodal region. This im-

plies that there exists an extinction threshold for the drive pressure amplitude, such that above this threshold a (sonoluminescing) bubble cannot be levitated *above* the pressure antinode. We note, however, that there does exist, at least theoretically, stable levitation positions *below* the pressure antinode above this extinction threshold. In this region, the Bjerknes force pushes the bubble down and away from the antinode, while the buoyancy force pulls the bubble up and toward the antinode, such that there can exist an equality between the Bjerknes and buoyancy forces.

Also note that the location of the zero-crossing of the Bjerknes force does not depend on the position of the bubble. This fact can be seen from Eq. (4). The zero-crossing shown in Fig. 3 depends only on the frequency and the equilibrium bubble size R_0 [and to a lesser extent, the surface tension and viscosity, as they relate to the radius $R(t)$]. Calculations over a large range of these values show only minor influences on the location of the zero-crossing. Experimentally, this extinction threshold does not appear to be reached. Instead, apparently shape instabilities are generated near a pressure amplitude of 1.5 bar, causing the bubble to selfdestruct.[17]

II. EQUILIBRIUM BUBBLE POSITION

In the analysis above, we assumed that the position of the bubble was 1 mm above the pressure antinode. This is due to the additional force of buoyancy exerted on a levitated bubble, directed vertically upward, against the gravitational field. This force is given by $F_{\mathrm{buoy}} = \rho g V(t)$, where ρ is the density of the fluid, and g is the gravitational acceleration. Here, we calculate the average levitation position of the bubble, and find how the equilibrium bubble position[18] is dependent on the drive pressure amplitude (assuming the equilibrium radius does not change).

A bubble can be stably levitated if the Bjerknes and average buoyancy forces are equal. Thus

$$\frac{\rho g}{T} \int_0^T V(t) dt = \frac{\omega k_z P_a}{T} \sin(k_z z) \int_0^T V(t) \sin(\omega t) dt. \quad (5)$$

A simple expression for the equilibrium levitation position of the bubble can be obtained provided one assumes the bubble is near the pressure antinode. Then $\sin(k_z z) \approx k_z z$, and

$$z \approx \frac{\rho g}{\omega k_z^2 P_a} \frac{\Lambda_1}{\Lambda_2}, \quad (6)$$

where $\Lambda_1 = \int_0^T V(t) dt$ and $\Lambda_2 = \int_0^T V(t) \sin(\omega t) dt$.

Figure 3 shows the equilibrium bubble position [using Eq. (6)] above the antinode as a function of the applied pressure amplitude. For small drive pressures, the equilibrium position of the bubble is nearly inversely related to the drive pressure, as seen in Eq. (6). However, as the drive pressure is increased further, the bubble position begins to shift *away* from the antinode, the slope of the curve changing signs. Beyond about 1.78 bar the bubble can no longer be levitated above the pressure antinode, since the Bjerknes force now pushes the bubble away from the antinode.

In the next section we describe our experimental apparatus and method for measuring the equilibrium bubble position, and attempt to confirm elements of the above analysis, specifically the region that shows the equilibrium position of
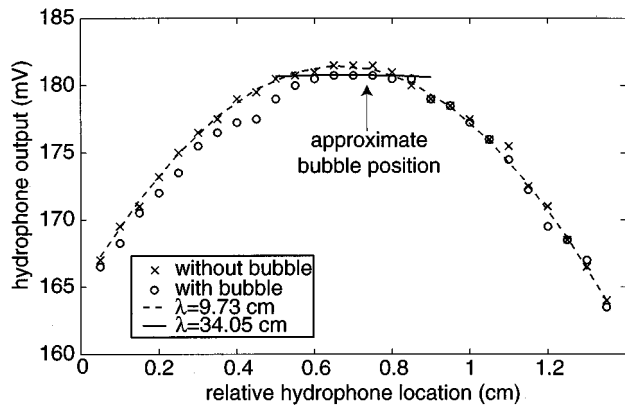
FIG. 4. The measured pressure profile in the levitation cell with ($\times$) and without ($\bigcirc$) a bubble present. The arrow marks the approximate location of the bubble during this measurement. Also shown is a fit to $A + B \cos(k_z z + \phi)$ for both cases. In the absence of the bubble, $\lambda = 9.73$ cm and corresponds to the third eigenmode of our system. With the bubble present, we set $\lambda' = 3.5 \times 9.73$ cm to take into account the possible flattening of the local pressure gradient, as discussed in the text. The particular value for λ' comes from determining the scaling factor that was needed to match the data with the calculations in Fig. 5. Note, however, that the variation between the two cases may simply involve acoustic radiation damping or scattering (the high-frequency acoustic signal emitted from the bubble was removed by passing the signal through a low-pass filter).

FIG. 5. Measurements of the levitation position of the bubble are shown together with the calculated results. The solid line comes directly from Eq. (6), using $\lambda = 9.73$ cm ($k_z \approx 64.6$ m^{-1}). The measurements, shown as circles, show qualitative agreement, but appear to be offset from the calculated values by a scaling factor. As discussed in the text, quantitative agreement can be obtained if one assumes that the bubble modifies the local pressure gradient, such that the bubble "sees" a wavelength of $\lambda' = 3.5\lambda$. The dashed curve represents a calculation that was modified to take this possible flattening of the pressure gradient into account. The inset shows a blow up of the measured region. The error bars are a result of bubble motion, as seen through the Gäertner microscope.

the bubble shifting further from the antinode as the drive pressure is increased. Parenthetically, we mention that in an experiment one cannot follow a bubble's position entirely along a single curve shown in Fig. 3. Each curve represents the equilibrium position for a particular bubble with a particular equilibrium size (3 and 5 μm). It is known that the equilibrium radius of a bubble changes as one changes the pressure amplitude.[17,19]

III. EXPERIMENT

Our experimental system was described above. It consisted of a rectangular levitation cell with dimensions 8.5 $\times$ 8.5-cm cross section, and a water column ($\rho = 1000$ kg/m^3) of 14.6 cm. The drive frequency for this system was 19.5 kHz, being operated in the (1,1,3) mode.

We measured the position of the bubble with a Sony Handycam video camera, focused at infinity, looking into a Gäertner microscope. Images were captured and stored on a computer for later processing. With filars of 10- and 100-μm separations (calibrated using a calibration microscope slide), resolutions as fine as 2 μm were obtainable. Backlighting from a fiber-optic illuminator was used as a (white) light source. Our results for the equilibrium position of the bubble will be shown in the next section.

The position of the sonoluminescence bubble was measured for different drive pressure amplitudes ($\approx 1.35 - 1.50 \pm 0.05$ bar). A calibrated needle hydrophone[20] was used to measure the approximate pressure amplitude near the antinode, with and without a bubble in place. Figure 4 shows the vertical pressure profile for the two cases. In both cases the hydrophone was offset approximately 4 mm from the center line of the cell, so that with the bubble present, we could still map out the vertical pressure profile in the cell, without disturbing the bubble.

The dashed line in Fig. 4 is a fit to $A + B \cos(k_z z + \phi)$ with $\lambda = 9.73$ cm. This corresponds to the vertical wavelength of our cell (recall that our cell is 14.6 cm long, and is operated in the third eigenmode). Note that with the bubble present, the pressure profile in the vicinity of the bubble varies slightly from that with no bubble present. The physical mechanism for this variation may be important, and is discussed in the next section.

IV. RESULTS AND DISCUSSION

Figure 5 shows the experimentally obtained equilibrium position of the bubble for various pressure amplitudes, together with the calculated equilibrium position. Note that there is good qualitative agreement between the calculated and measured equilibrium positions, showing that the equilibrium position of sonoluminescence bubbles actually shifts away from the pressure antinode as the drive pressure is increased. However, the quantitative agreement is poor. There appears to be a multiplicative scaling discrepancy of approximately 12.25. We now discuss possible sources for the observed discrepancy between the measured and calculated bubble position. (Parenthetically, we mention that in many experiments, one can generate sonoluminescing bubbles that move in almost any conceivable direction, and over relatively large distances, as the drive pressure is changed. Only after extreme care is taken to fine tune the driving frequency, does the bubble behave as shown here.)

Experimental errors include the fact that the absolute z coordinate of the bubble is estimated, since the exact location of the pressure antinode is only known as an average over the finite size of the hydrophone (≈ 1-mm diameter). Thus there may be a systematic offset to the bubble positions

measured with the video system. This error represents an error in the accuracy of the measurement; however, the relative error (error in precision) is small, since resolutions as fine as 2 μm were obtainable. In our case, the discrepancy between the measured and calculated positions appears to be that of scaling, and not from an offset error (note the log scale in Fig. 5).

Also, consider that in going from Eq. (5) to Eq. (6) we assume that the bubble is near the pressure antinode, and we take $\sin(k_z z) \approx k_z z$, where z is the distance above the antinode. Experimentally, the position of the bubble is $z' = z + c$, where c is an offset resulting from our inability to exactly determine the antinode position (due to the finite size of the hydrophone). Even an offset error as large as a few millimeters affects this approximation only minimally. A sensitivity analysis of Eq. (6) in other parameters such as the equilibrium radius of the bubble and drive pressure amplitude shows only small variations and do not account for the observed scaling discrepancy.

A potential experimental source for error that could result in a scaling error would be a miscalibration of the fiducial markings on the microscope eyepiece; however, these markings were calibrated against a calibration microscope slide. A video camera looking into the microscope eyepiece will not affect the calibration. Furthermore, a second set of experiments done at the fundamental frequency of 13.3 kHz (the 1,1,1 mode) also shows discrepancies, but with a different scaling (≈ 7.3).[21] One would expect that miscalibrated fiducial markings would result in similar scaling discrepancies.

It may be that the bubble itself, as an object in a sound field, provides a boundary which distorts the pressure profile in the cell, most profoundly near the bubble itself. Consider Fig. 4 which shows that the pressure profile with the bubble present appears flattened near the bubble (near the pressure antinode). This flattening of the pressure gradient may appear to the bubble as a longer localized wavelength. If so, then the bubble does not ''see'' a vertical wavelength of 9.73 cm, as calculated by the size of the water column in the levitation cell, but instead the flattening of the pressure gradient appears as a longer wavelength (at least in the near vicinity of the bubble).

If one assumes that the presence of the bubble affects the local gradient, resulting in a modified wavelength λ' $= 3.5\lambda$ [which becomes a scaling factor of 12.25 in Eq. (6), since $k_z = 2\pi/\lambda$], then quantitative agreement is obtained between the calculated and measured equilibrium position (as shown in Fig. 5). The introduction of a scaling factor via a modified wavelength is also self consistent with the measured pressure profile in the levitation cell near the bubble. Figure 4 shows that near the pressure antinode, the pressure profile can be fit with a wavelength of $\lambda' = 3.5\lambda$. Though not shown here, experiments done at 13.3 kHz also show quantitative agreement with the introduction of a modified wavelength of $\lambda' = 2.7\lambda$.

The analysis above did not take into account that a bubble in a stationary sound field can remove energy from the incident sound field through various mechanisms such as thermal, viscous, and radiative damping. For small bubbles driven into small amplitude oscillations, thermal damping dominates.[22] However, for large drive pressure amplitudes, radiative damping may dominate. These damping mechanisms may also show up as a change in the measured pressure profile. We were not able to conclusively determine whether the change in the pressure profile was due to damping or whether the bubble itself distorts the pressure gradient.

Nor did we consider the issue of bubble translation in this study. The translation may induce microstreaming to such an extent that it affects the bubble's equilibrium position. The issue of microstreaming from single-bubble sonoluminescence has only recently been considered.[23]

V. CONCLUSION

Measurements of the equilibrium position of a sonoluminescence bubble in a stationary wave shows qualitative agreement with the calculated equilibrium position; as the drive pressure amplitude increases, the bubble shifts away from the pressure antinode. We explained this shift as an equilibrium between the competing buoyancy and acoustic pressure forces. The delayed collapse of a sonoluminescence bubble at high-pressure amplitudes greatly affects the Bjerknes force. Furthermore, we found that at sufficiently high drive pressure amplitudes, there exists an upper pressure threshold for bubble levitation, even in the absence of surface instabilities. This threshold appears when the Bjerknes force no longer holds the bubble against the gravitational force (near 1.78 bar). At higher drive pressures, stable bubble levitation can only exist *below* the pressure antinode, where the Bjerknes and buoyancy forces again act in opposite directions (assuming the bubble remains spherical and does not develop other instabilities). These results may have implications for experiments performed in a microgravity environment, where the buoyancy force is negligible.

Quantitative agreement between the measured and calculated equilibrium positions of a sonoluminescence bubble can be obtained if one assumes that the bubble itself acts to distort the pressure profile in the immediate vicinity of the bubble. If true, then explicitly, how does the bubble affect the local pressure profile? Also, this may provide some insight into building stable levitation cells for single-bubble sonoluminescence experiments.

[1] D. F. Gaitan, L. A. Crum, C. C. Church, and R. A. Roy, ''Sonoluminescence and bubble dynamics for a single, stable, cavitation bubble,'' J. Acoust. Soc. Am. **91**, 3166–3183 (1992).

[2] B. P. Barber, R. Hiller, K. Arisaka, H. Fetterman, and S. Putterman, ''Resolving the picosecond characteristics of synchronous sonoluminescence,'' J. Acoust. Soc. Am. **91**, 3061–3063 (1992).

[3] M. J. Moran, R. E. Haigh, M. E. Lowry, D. R. Sweider, G. R. Adel, J. T. Carlson, S. D. Lewia, A. A. Atchley, D. F. Gaitan, and X. K. Marugama, ''Direct observations of single sonoluminescence pulses,'' Nucl. Instrum. Methods Phys. Res. B **96**, 651–656 (1995).

[4] V. H. Arakeri, ''Effect of dissolved gas content on single-bubble sonoluminescence,'' Pramana, J. Phys. **40**, L145 (1993).

[5] R. Hiller, K. Weninger, S. J. Putterman, and B. P. Barber, ''Effect of noble gas doping in single bubble sonoluminescence,'' Science **445**, 248–250 (1994).

[6] T. J. Matula, R. A. Roy, P. D. Mourad, W. B. McNamara III, and K. S. Suslick, ''Comparison of multibubble and single-bubble sonoluminescence spectra,'' Phys. Rev. Lett. **76**, 2602–2606 (1995).

[7] J. B. Young, T. Schmiedel, and W. Kang, ''Sonoluminescence in high magnetic fields,'' Phys. Rev. Lett. **77**, 4816 (1996).

[8] I. M. Hallaj, T. J. Matula, R. A. Roy, and L. A. Crum, ''Measurements of the acoustic emission from glowing bubbles,'' J. Acoust. Soc. Am. **100**, 2717(A) (1996).

[9] T. J. Asaki and P. L. Marston, ''Acoustic radiation force on a bubble driven above resonance,'' J. Acoust. Soc. Am. **96**, 3096 (1995).

[10] T. J. Asaki and P. L. Marston, ''Equilibrium shape of an acoustically levitated bubble driven above resonance,'' J. Acoust. Soc. Am. **97**, 2138 (1995).

[11] L. A. Crum and A. I. Eller, ''Motion of bubbles in a stationary wave,'' J. Acoust. Soc. Am. **48**, 181 (1969).

[12] L. A. Crum and A. Prosperetti, ''Nonlinear oscillations of gas bubbles in liquids: An interpretation of some experimental results,'' J. Acoust. Soc. Am. **73**, 121 (1983).

[13] V. Kamath, A. Prosperetti, and F. N. Egolfopoulos, ''A theoretical study of sonoluminescence,'' J. Acoust. Soc. Am. **94**, 248–260 (1993).

[14] A. Eller, ''Force on a bubble in a standing acoustic wave,'' J. Acoust. Soc. Am. **43**, 170 (1968).

[15] A. Prosperetti, ''Bubble phenomena in sound fields: part two,'' Ultrasonics **22**, 115 (1984).

[16] L. A. Crum, ''Bjerknes forces on bubbles in a stationary sound field,'' J. Acoust. Soc. Am. **57**, 1363 (1975).

[17] R. G. Holt and D. F. Gaitan, ''Observation of stability boundaries in the parameter space of single bubble sonoluminescence,'' Phys. Rev. Lett. **77**, 3791 (1996).

[18] The bubble actually translates vertically during an acoustic cycle. This is due to a change in the buoyancy and Bjerknes force with a change in the bubble's volume. In our work we simply average over an acoustic cycle, and do not consider the instantaneous position of the bubble.

[19] B. P. Barber, C. C. Wu, R. Löfstedt, P. H. Roberts, and S. J. Putterman, ''Sensitivity of sonoluminescence to experimental parameters,'' Phys. Rev. Lett. **72**, 1380 (1994).

[20] The hydrophone calibration is discussed in S. M. Cordry, ''Bjerknes forces and temperature effects in single-bubble sonoluminescence,'' Ph.D. thesis, University of Mississippi, 1995.

[21] S. M. Cordry, ''Bjerknes forces and temperature effects in single-bubble sonoluminescence,'' Ph.D. thesis, University of Mississippi, 1995.

[22] A. Prosperetti, ''Thermal effects and damping mechanisms in the forced radial oscillations of gas bubbles in liquids,'' J. Acoust. Soc. Am. **61**, 17 (1977).

[23] M. Longuet-Higins (private communication).

Measurements of coupled Rayleigh wave propagation in an elastic plate

Boon Wee Ti and William D. O'Brien, Jr.
Electrical and Computer Engineering Department, University of Illinois, Urbana, Illinois 61801

John G. Harris
Theoretical and Applied Mechanics, University of Illinois, Urbana, Illinois 61081

At frequencies where the thickness of an elastic plate is more than a wavelength, the propagation of the two lowest Rayleigh–Lamb modes in an elastic plate can be viewed as the propagation of a Rayleigh surface wave over two weakly coupled, surface-wave waveguides. That is, a Rayleigh wave launched on one surface gradually transfers to the other and then back. It does so in a length we call the beatlength. Measurements of the beatlength for brass plates are reported as a function of frequency and thickness. This phenomenon is readily excited and persists over a wide range of thicknesses and frequencies © *1997 Acoustical Society of America.*
[S0001-4966(97)05009-1]

INTRODUCTION

The two lowest waveguide modes of an elastic plate can be described as the propagation of a Rayleigh surface wave that starts on one surface, but that gradually transfers to the other surface. It then transfers back to the surface from which it started, the whole cycle taking place over a distance we call the beatlength. Provided the product of wave number and thickness is sufficiently large, one can thus view a plate as two weakly coupled, surface-wave waveguides. The objective of this letter is to report measurements of the beatlength in brass plates. This may be one of the first detailed experimental assessments of this phenomenon, though its possibility has been noted by Auld,[1] Viktorov,[2] and Brekhovskikh and Goncharov.[3] Moreover, Li and Thompson[4] made similar measurements, but did not widely report them.

The longer service life of structures such as pipelines means that they must be monitored more thoroughly, and over a longer period of time, for damage. Using coupled surface waves may be one way to inspect the inner, and therefore not easily accessible, surface of a pipe from its outer surface. Moreover, if the damage were a small surface-breaking fatigue crack, then a surface wave would readily detect the crack because the surface wave would strike the crack broadside, or if the damage were corrosion, then a surface wave would be more severely attenuated by the patch of corrosion at the surface than a bulk wave. The present measurements indicate that the coupled surface waves are readily excited and detected so that they can be used for such nondestructive testing.

The basic idea is presented using schematic drawings (Figs. 1 and 2). Figure 1 shows the dispersion relation for the two lowest Rayleigh–Lamb modes. The lower curve is that for the lowest antisymmetric mode, while the upper curve is that for the lowest symmetric mode. The vertical axis is the normalized angular frequency (ω multiplied by h/c_t, where h is one-half the thickness of the plate and c_t is the trans-verse wave speed). The horizontal axis is the normalized wave number (β multiplied by h). The diagonal short-dashed line indicates the straight line formed by $\omega h/c_t$ plotted against $\beta_r h$, where β_r is the Rayleigh wave number. The slope is c_r/c_t, where c_r is the Rayleigh wave speed. In the neighborhood of the intersection points of the horizontal, long-dashed line with the dispersion curves, the x particle displacements look roughly as sketched schematically in Fig. 2. If the symmetric and antisymmetric modes are both excited in phase, then the sum approximates a Rayleigh wave on the upper surface. This is indicated by the solid line in Fig. 2(c). However, each mode propagates with a slightly different wave number, β_s for the symmetric mode and β_a for the antisymmetric mode. After a distance $L/2$, the two modes move π out of phase. Adding the two modes together at this location approximates a Rayleigh wave on the lower surface. This is indicated by the dashed line in Fig. 2(c). After propagating an additional distance $L/2$, the modes move back into phase (more accurately 2π out of phase) and their sum again approximates a Rayleigh wave, now on the upper surface. In this sense the propagation of the two modes can be viewed as a Rayleigh surface wave coupling from one surface-wave waveguide to another. Figure 1, by means of the vertical dashed lines, indicates the difference 2ε between the wave numbers β_s and β_a. The difference between β_s and β_r is almost equal to that between β_r and β_a. The beatlength L of the coupled waves is that distance over which the two modes move out of phase by 2π, that is,

$$(\beta_a - \beta_s)L = 2\pi \tag{1}$$

or

$$L/h = \pi/h\varepsilon. \tag{2}$$

With Eq. (5.15), Brekhovskikh and Goncharov[3] give an estimate of ε for large βh. However, we use the exact Rayleigh–Lamb dispersion relation[5] to calculate ε and L/h.
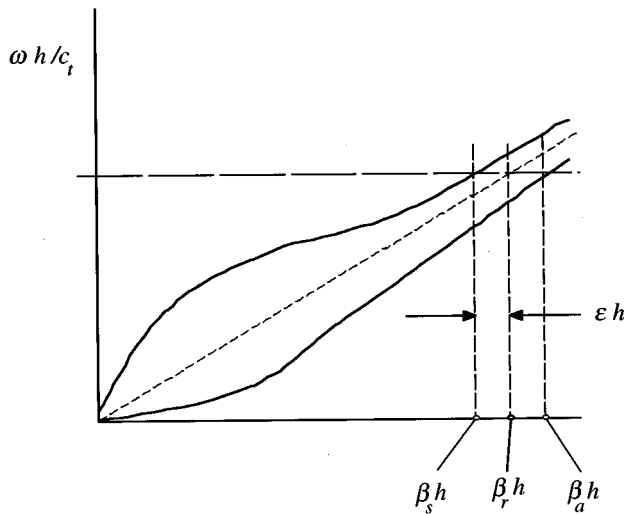
FIG. 1. A sketch of the dispersion relation for the two lowest modes of an elastic plate. The upper curve is the symmetric mode and the lower is the antisymmetric mode. The beatlength $L = \pi/\varepsilon$. Figure 2 sketches the particle displacement in the neighborhood of the intersections of the long-dashed horizontal line with the dispersion curves.

FIG. 2. A sketch of the x particle displacements corresponding to the neighborhoods of the intersection of the long-dashed horizontal line in Fig. 1 with each dispersion curve. The solid lines indicate that the modes are in phase, while the dashed lines indicate that they are π out of phase: (a) symmetric mode, (b) antisymmetric mode, and (c) algebraic sum of the two modes (schematic only).

I. MEASUREMENTS

The longitudinal and shear wave speeds in brass are experimentally determined using standard pulse-echo techniques with a Panametrics 5800 Pulser/Receiver. From the acquired front and back wall echoes, the time difference between these echoes was calculated with a correlation procedure and, using the measured plate thickness, the wave speed was calculated.[6] The longitudinal wave speed is found using a 15-MHz Panametrics V319 immersion transducer. The shear wave speed is found using a 20-MHz Panametrics V222-BA-RM normal incidence, shear-wave, contact transducer.

The view, from above, of the experimental arrangement is sketched in Fig. 3. A focused 3-MHz Panametrics V3680 transducer directs a focused ultrasonic beam at the surface of a 2.38-mm-thick brass plate at the Rayleigh angle θ_r. Because we were imitating a possible nondestructive testing measurement we did not measure the plate's thickness systematically, using a micrometer (accuracy about ± 13 μm) to measure the thickness at only one or two locations. At 3 MHz the longitudinal wavelength, in the brass plate, is approximately 1.5 mm and the shear wavelength approximately one-half that. The coupling phenomena we were seeking would be of little use in nondestructive testing if it were sensitive to variations in thickness on the order of 1/100 of a wavelength.

The plate is partially submerged in water. The portion submerged in water is used to position the transducer relative to the plate surface via pulse-echo, normal incidence, time-of-flight measurements. That above water is used to make the beatlength measurements. There the sound is coupled to the plate by a jacket of water that surrounds the sound beam, as shown in Fig. 3. The portion of the plate, in air, experiences no substantial fluid loading. The same function generator was used for the positioning measurements and the beat-

length measurements, with only the gate length being changed.

The normal beam axis to the brass plate is found by adjusting the focused transducer to an angle where the reflection, measured using the pulse-echo mode, is maximum. The parallelism of the plate to the scanning motion of the transducer was adjusted by using the same pulse-echo measurement to measure the distance at a few randomly selected points, on that part of the plate submerged in water, to make the distances equal. The precision positioning system is computer controlled and has a linear accuracy of about 5 μm. Assuming that the parallelism was maintained to $\pm\frac{1}{4}$ of a wavelength, in water at 3 MHz, the accuracy would be ± 0.12 mm. The transducer is then rotated to the Rayleigh angle θ_r. This angle is calculated (see Sec. III) from the measured wave speeds. The rotational accuracy of the precision positioning system, also computer controlled, is about $0.02°$.

The beatlength is experimentally determined as follows. With the incident beam at the Rayleigh angle and the focal point on the metal plate, the transmitter is moved parallel to the plate surface, toward the broadband receiver, a Deci model SE 1025-H308 surface contact transducer. The area of the intersection of the beam with the plate's surface is not
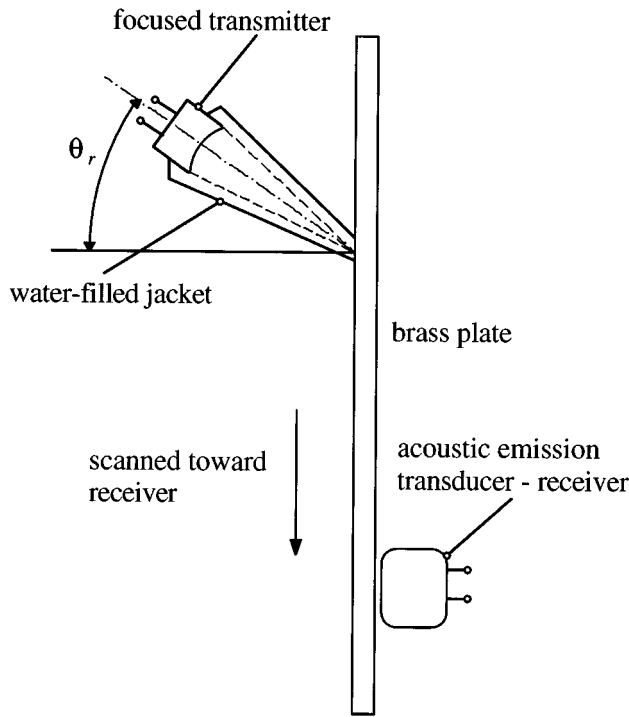
FIG. 3. A drawing of the experimental arrangement viewed from above. The central axis of the focused transducer makes an angle θ_r with the vertical. The sound is coupled to the plate by means of a water-filled jacket. Though the lower part of the plate is in water (this is not shown), the part used for the beatlength measurement is loaded only by the surrounding air. The focal point is placed at the plate surface and moved toward the stationary receiving transducer.

FIG. 4. Representative data, case 4 of Table I. (a) The spatially amplitude-modulated, received signal plotted against position along the plate's surface (500 rms data points, 100-mm scan distance). (b) The power spectrum of the signal from (a) plotted against the spatial frequency (1/distance). The vertical line indicates the theoretically estimated, spatial beat frequency.

II. RESULTS

The measured longitudinal wave speed in brass is 4536 m/s. The measured shear wave speed in brass is 2215 m/s. Their ratio is 2.048. The measured wave speeds had an accuracy of about ±0.1%.

Figure 4 shows a typical measurement (case 4 of Table I). Figure 4(a) shows a single realization of the amplitude modulated signal from a 100-mm scan, that is, the transmitter was moved 100 mm, in 200-μm intervals, toward the receiver. Figure 4(b) shows the spatial power spectrum of the signal from Fig. 4(a). The higher spatial frequency components evident in Fig. 4(a) are outside of the spatial frequency range of Fig. 4(b). The vertical line marks the theoretically predicted spatial beat frequency calculated using the exact Rayleigh–Lamb dispersion relation.[5]

Figure 5 plots L/h against $\omega h/c_t$. The error bars represent ±10% of a theoretically predicted value. The transmitter's operating frequency f is varied to produce the seven different $\omega h/c_t$ cases (Table I).

Table I summarizes the results of the experiment conducted on brass. The wave number β must be greater than ε

important beyond the need to keep it small so that the plate is not fluid loaded. The transmitter is moved in 200 μm intervals, and at each interval, a 10-MHz A/D rate, 1024-point data record of the temporal, received sinusoidal signal is recorded by the receiver. The rms value for each position is calculated and recorded.

The transmitter is moved a distance slightly greater than three times the estimated beatlength. The rms values for each position are Fourier transformed without windowing using a 4096 FFT. The power spectrum is calculated to determine the peak corresponding to the beatlength.

TABLE I. The measured and calculated values of the important parameters. The temporal frequency f is the transducer operating frequency, f_a is the spatial frequency, L is the beatlength, and h is the half-thickness of the brass plate. Note that case 4 is illustrated in Fig. 4.

				Theory			Experiment		
Case	f (MHz)	$\omega h/c_t$	β/ε	f_a (m^{-1})	L (mm)	L/h	f_a (m^{-1})	L (mm)	L/h
1	0.733	2.59	6.88	109.8	9.1	7.6	97.7	10.2	8.6
2	0.800	2.82	7.64	88.9	11.3	9.4	80.6	12.4	10.4
3	0.866	3.06	13.01	72.3	13.8	11.6	63.5	15.7	13.2
4	0.933	3.29	17.14	59.7	16.7	14.1	58.6	17.1	14.4
5	1.00	3.53	22.11	50.0	20.0	16.8	45.2	22.1	18.6
6	1.20	4.23	43.73	30.8	32.5	27.3	28.1	35.6	29.9
7	1.33	4.70	66.03	22.9	43.7	36.7	22.0	45.5	38.2
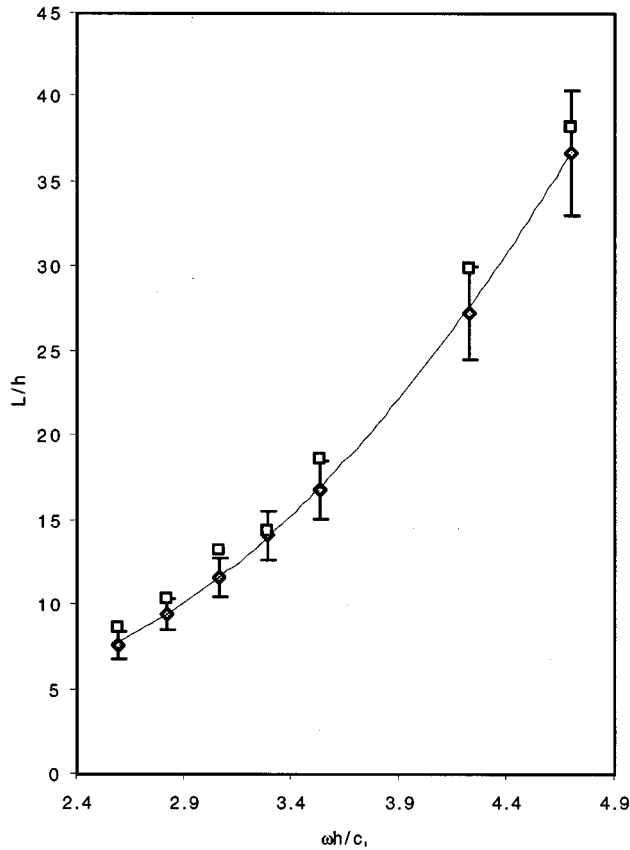
FIG. 5. A plot of L/h against $\omega h/c_t$. against The data points indicated by open squares are the measured results. The data points indicated by open diamonds are the theoretically predicted results. A $\pm 10\%$ error bar is attached to the theoretically estimated results.

to produce a spatially modulated signal. Thus β/ε is recorded. The theoretical value for L/h is calculated from the exact dispersion relations. The experimentally determined spatial beat frequency, beatlength, and normalized beatlength are tabulated in the last three columns. While earlier we indicated, to some extent, the accuracy of our measurements, we have not done so in this table because the measurement we sought was not sensitive to these inaccuracies.

III. DISCUSSION

The Rayleigh angle, θ_r, is defined as

$$\theta_r = \sin^{-1}(c/c_r), \tag{3}$$

where c is the speed of sound in water ($c = 1482$ m/s at 20 °C). A very good approximation for c_r is[7]

$$c_r/c_t = (0.87 + 1.12\nu)/(1 + \nu), \tag{4}$$

where

$$\nu = [(c_l/c_t)^2 - 2]/2[(c_l/c_t)^2 - 1]. \tag{5}$$

From the measured values of c_l and c_t, Eq. (4) gives $c_r = 2069$ m/s and Eq. (3) gives $\theta_r = 45.78°$. This is the angle used in these measurements.

Figure 4(b) suggests that the beat phenomenon is robust in the spatial frequency domain, and that the measured spatial frequency is in agreement with theory (see Table I). Figure 5 indicates that the measured values of the spatial beat frequency lie close to or within $\pm 10\%$ of the theoretically predicted values. Brekhovskikh and Goncharov[3] suggest that βh must be large for the coupling phenomena to be clearly observed and their estimate of ε assumes this. However, our values of βh ranged from 2 to 5, and yet we observe the coupling phenomena without difficulty. Moreover, the beatlengths indicate that the use of the coupled waves to access an inner surface is realistic. None of the beatlengths is so great that the signals would become too severely attenuated, with distance, to carry information from the far surface.

Initially, we had been concerned that, in addition to the two lowest modes, we might excite higher modes and as a consequence some power would be carried by them and lost to the coupling phenomenon. This appears not to be the case suggesting that when we are incident at the Rayleigh angle we strongly excite only the two lowest modes.

ACKNOWLEDGMENTS

[1] B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990), Vol. 2, pp. 93–94.

[2] I. A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967), pp. 93–96.

[3] L. Brekhovskikh and V. Goncharov, *Mechanics of Continua and Wave Dynamics* (Springer-Verlag, New York, 1985), pp. 75–81. Note that the first of Eqs. (5.16) has a sign error. The plus sign between the two terms should be replaced by a minus sign.

[4] Private communication, R. B. Thompson, Iowa State University.

[5] B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990), Vol. 2, pp. 77–82, Eqs. (10.18) and (10.19).

[6] I. A. Hein and W. D. O'Brien, Jr., "A real-time ultrasound time-domain correlation blood flowmeter, part II, performance and experimental verification," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **40**, 776–785 (1993).

[7] I. A. Viktorov, *Rayleigh and Lamb Waves* (Plenum, New York, 1967), p. 3.

Direct experimental investigations of acoustic modes guided by a solid–solid interface using optical interferometry

Ch. Matteï,[a)] X. Jia, and G. Quentin

Groupe de Physique des Solides, Université Paris 7, Unité associée au CNRS No. 17, tour 23,
2 place Jussieu, 75251 Paris Cedex 05, France

This paper presents direct field measurements of acoustic modes guided by the interface between two transparent solids. The measurement technique is based on the acousto-optical interaction inside the solid between the acoustic field and the probe laser beam of an interferometer. The main advantage of the method is its ability to measure acoustic strain fields in areas of difficult access with the classic detection methods. Moreover, it gives complete information about the dilatation strain field inside the solid, e.g., amplitude and phase. The propagation of a real velocity mode (Stoneley wave) is first illustrated. Then the situation of complex velocity modes is investigated for a Plexiglas–fused quartz slip interface. This material combination supports two possible interface modes theoretically. These modes are simultaneously observed and the differences between their behavior are measured. © *1997 Acoustical Society of America.* [S0001-4966(97)02006-7]

PACS numbers: 43.35.Sx, 43.35.Yb [HEB]

INTRODUCTION

Characterization of elastic properties of interfaces between solid media is a subject of fundamental interest in such fields as nondestructive testing[1] or geophysics.[2] Acoustic bulk waves, e.g., longitudinal or transverse waves, are sensitive to discontinuities that are characteristic of bond failures. Several methods[3,4] have been developed in order to obtain information on interface qualities from transmission or reflection measurements of these acoustic bulk waves but these methods are hardly sensitive to fine details of interface properties. On the other hand, acoustic guided modes are confined in the interface region[5,6] and their propagation properties are much more influenced by interface characteristics.[7] Even so, only limited results of direct measurements of these guided waves have been obtained because of the difficulties associated to the interface access. As a matter of fact, most of these experimental works are phase velocity measurements using classic piezoelectric transducers. Lee and Corbly[8] presented measurements of Stoneley waves and attenuated interface waves velocities in the aim of interface inspection.

An alternative to the classic piezoelectric transducer methods is the detection of interface waves by optical means. Claus and Palmer[9] showed that optical detection methods could give more information than classical measurements; they detected the normal displacement of the interface and measured the Stoneley wavelength at a nickel–Pyrex interface.

Recent developments in optical interferometers dedicated to ultrasonic measurements made possible new types of measurements. Recently, Jia *et al.*[11] described a method based on the acousto-optic interaction which permits a local measurement of the dilatation induced by a guided acoustic wave propagating inside a transparent solid. The main advantage of the method is its ability to measure acoustic strain fields in areas uneasy to probe with the classic detection methods. Moreover, it gives complete information on the dilatation strain field inside the solid, e.g., amplitude and phase.

In the context of interface waves studies, direct measurement of the acoustic field close to the interface is of primary interest. Measuring directly the inhomogeneity of the dilatation strain field gives a straightforward determination of the interface wave nature and it is the aim of this paper to present direct measurements of different kinds of interface waves using the interferometric detection method. Specifically, the simultaneous propagation of two leaky interface waves is presented. To our knowledge, this situation, which was predicted by Pilant[12] in a study of the complex roots of the solid–solid interface characteristic equation, has never been experimentally observed.

In the first part of the paper, principles of the interferometric detection method are presented and the experimental setup is shown. In the second part the classical characteristic equation roots problem for an interface between two solids is briefly presented and the nature of different interface modes is illustrated. The third part contains experimental results obtained with the interferometric detection method. First, direct measurements of the ''Stoneley'' wave (i.e., an interface wave propagating without loss along the interface) are shown for two different boundary conditions and two combinations of materials. Then the case of the simultaneous propagation of two leaky interface waves is presented. Velocities, attenuation, and field inhomogeneity are measured for both of them. The characteristics of these two interface modes are compared with the theoretical predictions and the physical difference between these two modes is discussed.

[a)]Present address: Welding Engineering Institute, Ohio State University, 190 W. 19th Avenue, Columbus, OH 43210.
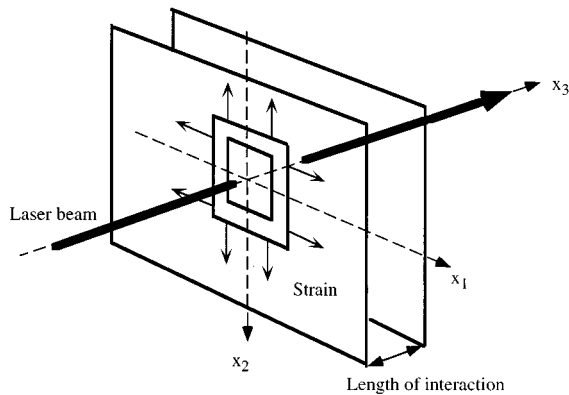
FIG. 1. Geometry of the acousto-optical interaction used for the interferometric detection method.

I. OPTICAL DETECTION TECHNIQUE AND EXPERIMENTAL SETUP

The optical detection technique used to investigate interface waves is based on an interferometric measurement of the dilatation induced by an acoustic guided wave propagating inside transparent materials. Like most of the acousto-optical detection techniques, this measurement is limited to two-dimensional (2-D) acoustic field probing. This technique has been applied first to Rayleigh wave detection[13] and to Lamb wave-field measurements inside transparent plates.[11] This section presents the principles of the detection method. A more complete analysis of the acousto-optic interaction involved in this technique has been presented in a previous paper.[11]

When an acoustic wave propagates inside an isotropic transparent solid of optical index $n = \epsilon^2$, the solid becomes optically birefringent and a light beam crossing the acoustic wave orthogonally undergoes a phase shift. Let us consider a guided acoustic wave with its displacement components in the sagittal plan (O, χ_1, χ_2), defined by its strain tensor S_{kl} as shown in Fig. 1. As a result of the acoustic pertubation, the solid medium becomes optically anisotropic. The corresponding changes for the dielectric tensor are[14,15]

$$\Delta \epsilon_{ij} = \Delta(n^2)_{ij} = -\epsilon^2 p_{ijkl} S_{kl}, \qquad (1)$$

where n is the optical index of the medium and p_{ijkl} are the photoelastic constants of the material. The effect of the optical index variations on a beam crossing the solid medium are twofold: the transmitted light exhibits polarization and phase changes.[15] For a laser beam much narrower than the acoustic wavelength, crossing an acoustic beam orthogonally propagating in the x_3 direction, the variation of the optical index induces a phase shift $\delta\varphi$:[11]

$$\delta\varphi = -n^3 \frac{\omega l}{2c} (p_{11} + p_{22})(S_1 + S_2), \qquad (2)$$

where ω is the optical frequency, c the velocity of the light in the material, and l the lateral dimension (measured along χ_3) of the acoustic beam and where the photoelastic constants and strains are written using the usual reduced matrix notation.
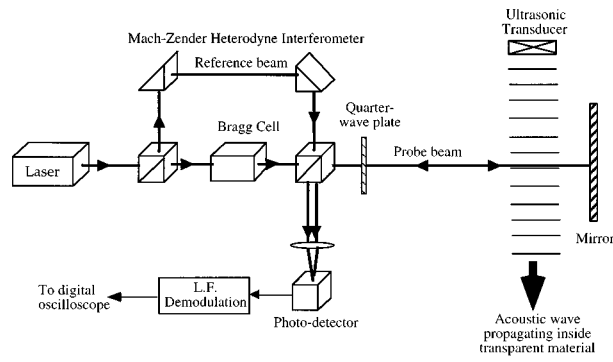
FIG. 2. Schematic diagram of the dilatation strain measurement with a Mach–Zender interferometer.

This optical phase shift can be measured using a Mach–Zender interferometer.[16] The light coming from a He–Ne laser is divided into two beams (Fig. 2). The reference beam goes directly to a photodiode, whereas the probe beam, shifted in frequency by a Bragg cell ($f_b = 70$ MHz), crosses the acoustic beam back and forth and interferes with the reference beam on a photodiode. The phase shift $\delta\varphi$ is demodulated from the photocurrent by a broadband electronic processing used in acoustic displacement measurements.[16] Finally, the signal at the output of the interferometer is a time signal, proportional to the acoustic dilatation strain, containing both amplitude and phase information. A sensitivity of 1 mV for a 10-7 strain is obtained inside fused quartz.[11]

The quantity measured ($S_1 + S_2$) is analog to the relative dilatation $\Delta V/V$ inside the solid and will be noted underbar $\underline{\Delta}$. It can be expressed for harmonic acoustic waves as

$$\underline{\Delta} = (S_1 + S_2) = \operatorname{div} \mathbf{u} = k^2 \phi, \qquad (3)$$

where $\mathbf{u}$ is the particle displacement vector, k the wave number, and ϕ the longitudinal potential.

The experimental setup is shown in Fig. 3. The samples are placed between the interferometer and a steady mirror. The laser beam orthogonally crosses back and forth the transparent materials to the acoustic beam and the optical phase shift induced by the optical index variations is detected by the interferometer. The corresponding electrical signal is
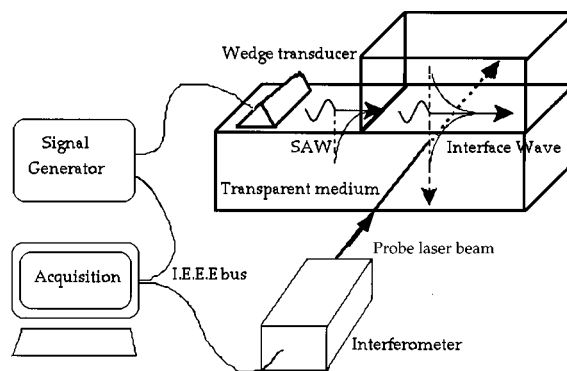
FIG. 3. Experimental setup used to generate and detect interface waves: The interface wave is generated by the conversion of a Rayleigh wave. The laser beam coming from the interferometer is crossing perpendicularly the transparent material.
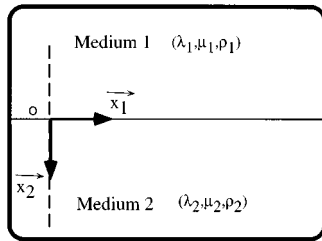
FIG. 4. Coordinate system for interface waves propagation.

collected at the output of the interferometer and sampled on a digital oscilloscope. Signal processing and representation is done on the sampled signal. By moving the sample in front of the interferometer, scans in the x_1 or x_2 direction can be done allowing velocity and attenuation measurements in the propagation direction as well as amplitude measurement in the direction normal to the interface.

II. INTERFACE WAVE PROPAGATION

In this section, we present briefly the solid–solid interface problem. The Cartesian system used in this description is shown in Fig. 4. The subscript 1 is dedicated to the upper medium and subscript 2 to the lower medium. The materials are defined by their Lamé constants and densities, λ_1, μ_1, ρ_1 and λ_2, μ_2, ρ_2, respectively. In this system, the solutions of the equation of linear elasticity in each material can be written in terms of the longitudinal Φ and shear Ψ potentials in each medium,

$$\Phi_1 = A_1 \exp(ik_z^{l_1}z)\exp(i(k_xx-\omega t)),$$

$$\Psi_1 = B_1 \exp(ik_z^{t_1}z)\exp(i(k_xx-\omega t)),$$

$$\Phi_2 = A_2 \exp(-ik_z^{l_2}z)\exp(i(k_xx-\omega t)),$$ (4)

$$\Psi_2 = B_2 \exp(-ik_z^{t_2}z)\exp(i(k_xx-\omega t)),$$

where $k_z^{l_i} = \sqrt{k_{l_i}^2 - k_x^2}$ and $k_z^{t_i} = \sqrt{k_{t_i}^2 - k_x^2}$ with $i=1,2$, and where k_x is the wave vector of the guided wave along the interface, and k_{l_i}, k_{t_i} are the longitudinal and shear wave number in the medium (i).

Two different sets of boundary conditions corresponding to experimental situations are chosen. The first ones are the slip boundary conditions, i.e., continuity of normal displacement, continuity of normal strain, and cancellation of transverse stress:

$$u_2(x_2=0^+)=u_2(x_2=0^-),$$

$$T_{22}(x_2=0^+)=T_{22}(x_2=0^-),$$ (5)

$$T_{12}(x_2=0^+)=0, \quad T_{12}(x_2=0^-)=0.$$

The second set describes the classic ''bonded'' conditions, i.e., continuity of both displacements and stresses:

$$u_1(x_2=0^+)=u_1(x_2=0^-),$$

$$u_2(x_2=0^+)=u_2(x_2=0^-),$$

$$T_{22}(x_2=0^+)=T_{22}(x_2=0^-),$$ (6)

$$T_{12}(x_2=0^+)=T_{12}(x_2=0^-).$$

In each case, writing the four boundary conditions as a function of the potentials leads to a system of four linear equations with nontrivial solutions when its determinant is equal to zero.

The existence domain of a lossless interface mode (i.e., a real root, k_x, of the characteristic equations corresponding to the so-called Stoneley wave) has been studied by several authors. Scholte[6] first showed that only a narrow range of material combinations support such interface waves in the case of bonded boundary conditions. Murty[7] showed that this range was larger in the case of slip boundary conditions.

When the characteristic equation admits only complex values of k_x as solutions, Pilant[12] showed that, depending on the materials combination, two roots corresponding to leaky interface waves, i.e., roots with small imaginary parts, exist. He showed numerically that when the shear modulus of one of the material tends to zero (fluid–solid interface), one of these solutions approaches asymptotically the generalized Rayleigh wave. Under the same conditions, the other solution approaches the Scholte–Stoneley wave. By extension, he called the first mode the Rayleigh wave, and the second the Interface wave. Moreover, Pilant studied the existence domain for these two interface waves in function of the materials elastic constant in the case of bonded boundary conditions and found a range of material where both the Interface and Rayleigh waves can propagate. To our knowledge, the simultaneous propagation of these two modes has never been experimentally confirmed.

These three interface waves, Stoneley wave, Rayleigh wave, and Interface wave, have different field inhomogenities which can be determined from the calculation of the potentials or the displacement fields in each material. The description of these interface modes in terms of a combination of inhomogeneous waves is also another way to determine these mode's inhomogeneities.[17] However, these two methods show the physical differences between these interface modes. As an example, the variation of the potentials in the direction perpendicular to the interface in each material is illustrated in Fig. 5 for the three interface waves. For the Stoneley wave, maximum amplitude for each potential is located near the interface. The two other modes are different in nature because at least one potential increases with increasing distance to the interface, i.e., part of the energy reradiates in one medium. Moreover, the radiation mechanism is not the same for these two modes. In the case of the Rayleigh wave, both longitudinal and shear potential amplitudes increase with the distance to the interface in one medium, whereas for the Interface mode the longitudinal potential is confined at the interface.

In practice, one of the difficulties when dealing with interface wave measurements is to clearly identify the measured interface modes among the 16 roots of the characteristic equation. Comparison between the measured and calcu-
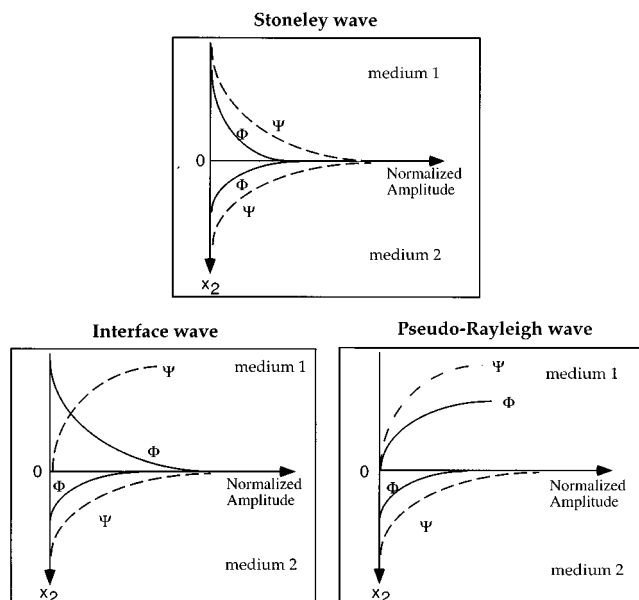
FIG. 5. Variation of the longitudinal (Φ) and shear (Ψ) potentials with the distance to the interface for the three interface waves involved: The Stoneley wave, the Interface wave, and the pseudo-Rayleigh wave (after Pilant).

lated velocities is the only way to achieve this identification with the classic piezoelectric detection techniques. The main feature of the interferometric detection method is its ability to measure the dilatation strain field, related to the longitudinal potential, in the materials as long as they are transparent. It gives the opportunity to measure in a simple way the field inhomogeneities of these interface modes. This information can be helpful in order to identify and characterize each mode as illustrated in the following section.

III. DIRECT MEASUREMENT OF INTERFACE WAVES

This section presents the experimental results obtained with the interferometric detection method. Three different combinations of materials are selected. In order to detect the Stoneley wave, the tungsten–fused quartz (bonded and slip boundary conditions) and the fused quartz–fused quartz (slip boundary conditions) combinations were selected. The case of leaky interface waves is investigated for the fused quartz–Plexiglas slip interface. The acoustical parameters of the three materials involved in these combinations are given in Table I.

The slip contact was achieved by setting a thin layer of water in between the optically polished materials. The thickness of the layer, estimated to $1–2$ μm for water is far

TABLE I. Elastic characteristics of the different materials used in the experiments.

	Longitudinal velocity (m/s)	Transverse velocity (m/s)	Density kg/m^3
Fused quartz	5960	3760	2200
Tungsten	5220	2890	19 300
Plexiglas	2680	1100	1800

smaller than all the acoustic wavelengths. In the case of bonded contact, a film of glue (with a thickness estimated to 10 μm) is set in between the materials. The interface waves were generated using the conversion of a surface wave (Rayleigh wave) at the solid–solid interface. The Rayleigh wave is generated using the classic wedge technique, as shown in Fig. 2. Broadband pulses with a 1 MHz central frequency are used to generate the surface wave.

A. Direct Stoneley wave measurement

The characteristics of the detection technique are first illustrated in the case of a tungsten–fused quartz interface which supports a Stoneley wave for both bonded and slip boundary conditions. Because tungsten is a nontransparent material, the strain field measurements are only possible on one side of the interface, i.e., inside fused quartz. Figure 6 shows two records of the Stoneley wave taken at two different depths from the interface in the case of a bonded and a slip interface. The dilatation strain field confinement inside fused quartz is clearly seen in both cases.

By scanning the field in the direction parallel to the interface, velocity and attenuation measurements can be achieved. The measured velocities are 2.73×10^3 (± 30) m/s for the slip interface and 2.86×10^3 (± 30) m/s in the case of the bonded interface, in good agreement with the theoretical ones, respectively, 2.78×10^3 m/s and 2.89×10^3 m/s. In order to make sure that no attenuation occurs during the propagation, the spectra of two pulses are displayed on the same graph (Fig. 7). No evidence of attenuation is seen in both cases.

In order to scan the dilatation strain field on each side of the interface, a Stoneley wave is generated along a fused quartz–fused quartz slip interface. Figure 8 shows four time signals corresponding to a scan in the direction perpendicular to the interface. Only one mode is observed and this mode is clearly confined at the interface. The dilatation field presents a π phase shift on each side of the interface due to the flexural characteristics of the boundary motion. A scan in the direction parallel to the interface shows that the pulse propagates without measurable attenuation along a distance of 10 cm. By measuring the time of flight, a velocity of 3.34 $\times 10^3$ (± 30) m s^{-1} is observed, close to the theoretical velocity of the Stoneley wave (3.36×10^3 m/s).

To measure quantitatively the decays in the x_2 direction of this mode, a method based on the Fourier spectral analysis of two records corresponding to two different depths is performed. In the case of a inhomogeneous plane wave, the Fourier transform of a dilatation strain pulse can be written without loss of generality

$$\Delta(x_1, x_2, k) = A(k) f(x_1) \exp(-\beta k x_2). \qquad (7)$$

The term $A(k)$ is characteristic to the generation of the pulse and $f(x_1)$ to the propagation. Figure 9(a) shows two spectra of time signals corresponding to two measurements of the same pulse measured at constant x_1 but for two different x_2 in the medium 1. The ratio of these two spectra is

$$\exp(-\beta k(x_2' - x_2)), \qquad (8)$$

Slip boundary conditions

Bonded boundary conditions
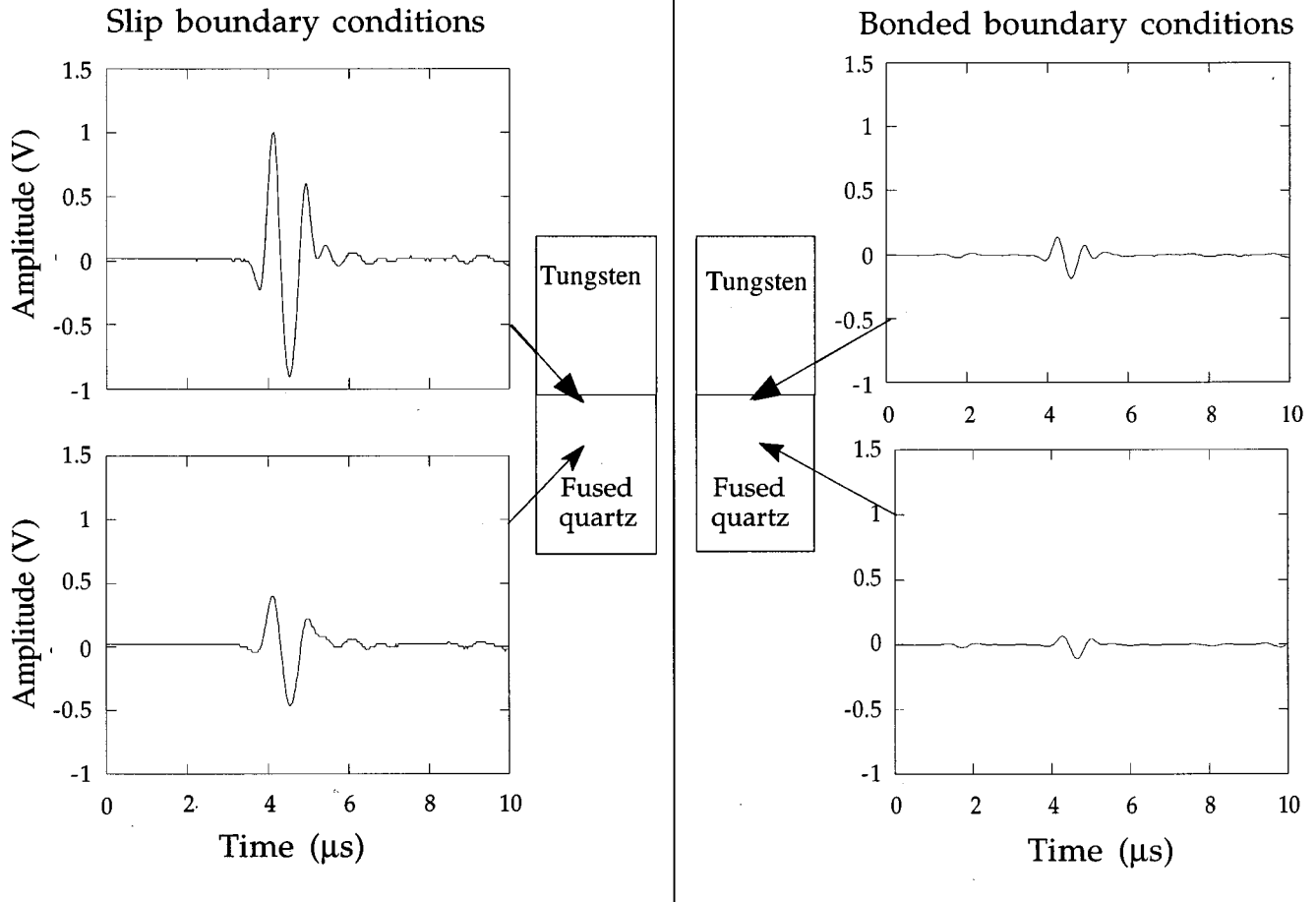
FIG. 6. Time signals corresponding to a scan in the z direction inside fused quartz for the tungtsen–fused quartz for the bonded interface and the slip interface cases.

where x_2' and x_2 are the two absolute positions of the records. The same procedure can be applied in the other medium.

Figure 9(b) shows a plot of the measured decay on one side of the interface represented as a function of the fre-
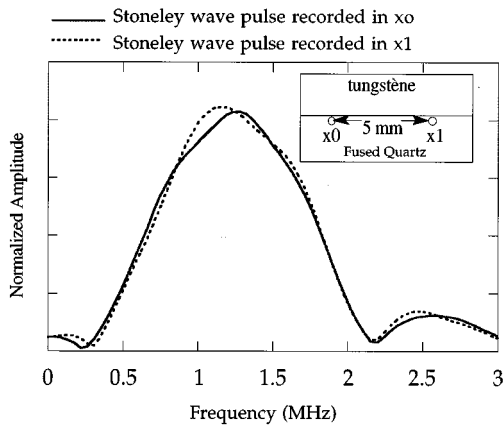
FIG. 7. Spectra of two time signals of a Stoneley wave at the tungsten–fused quartz interface recorded at the same depth and for two positions along the x_1 axes separated by 5 mm. No attenuation during the propagation is seen.
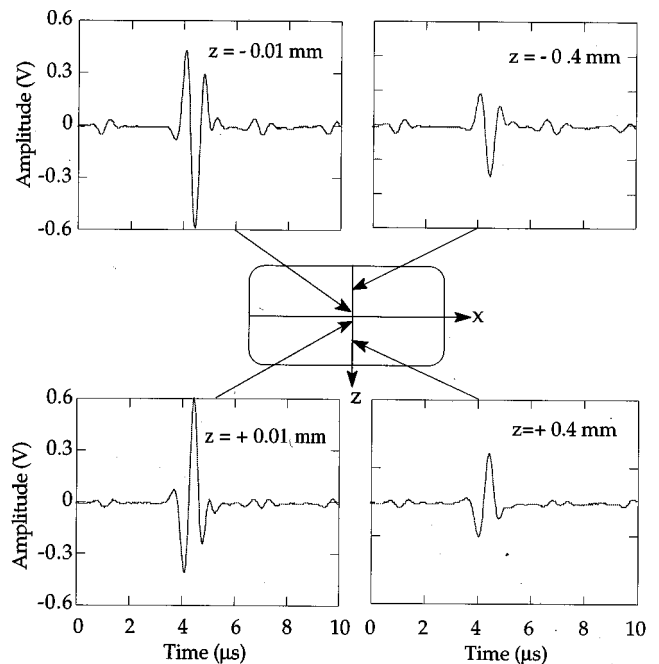
FIG. 8. Time signals corresponding to a scan in x_2 direction for the fused quartz–fused quartz slip interface.

FIG. 10. Time signals corresponding to a scan in the x direction for the Plexiglas–fused quartz slip interface. The Rayleigh wave (R) and the interface wave (I) are clearly seen.

FIG. 9. Spectra of the time signals recorded for two different depths (a) and their ratio (b) showing the Stoneley wave decay in the x_2 direction as a function of frequency. The dashed line in (b) shows the calculated dilatational strain for the corresponding experimental conditions.

quency. The measurements are in good agreement with the predicted inhomogeneity of the Stoneley wave for this material combination.

B. Direct measurement of leaky interface wave

The case of a Plexiglas–fused quartz slip interface has been investigated. This combination of material supports do not support a Stoneley wave. If a interface wave propagates, it should be either an interface wave or a Rayleigh wave as described previously. As a matter of fact, a numerical investigation of the characteristic equation shows the existence of two complex roots corresponding to interface waves which can be excited:

$$k_x = (1.079 + i0.086)k_{t_{\text{fused quartz}}},$$
$$k_x = (1.350 + i0.003)k_{t_{\text{fused quartz}}}, \tag{9}$$

where $k_{t_{\text{fused quartz}}}$ is the shear wave number inside fused quartz. This material combination supports the simultaneous propagation of a Rayleigh wave [first root in (9)] and an interface wave and the aim of this part is to illustrate this special situation and to characterize the physical difference

between these two modes using the interferometric detection method.

Figure 10 shows three time signals corresponding to a scan in the direction of propagation taken in the lower medium, i.e., fused quartz. The first signal is taken just after the conversion of the initial Rayleigh wave. Then during the propagation, two modes propagating with different velocities are clearly seen. In addition, these two modes propagate with different losses.

In order to characterize these two modes, the comparison between their velocities, attenuations, and field confinement at the interface is achieved. The measured velocities are 3.50×10^3 m/s for the faster mode and 2.70×10^3 m/s for the slower mode. The first mode propagates with a velocity closed to the Rayleigh velocity in fused quartz (3.36×10^3 m/s), the second propagates with a velocity close to the longitudinal velocity in Plexiglas (2.68×10^3 m/s). These velocities are close to the ones deduced from the real part of the two solutions (9) which are 3.48×10^3 m/s and 2.71×10^3 m/s.

In order to measure attenuation along the direction of propagation, the principle of the procedure previously described for the inhomogeneity in the direction normal to the interface is applied, but the spectrum analysis is achieved for each pulse measured at constant x_2 and two different x_1.

The corresponding losses for each modes are plotted as a function of the adimensional quantity $\Delta x / \lambda_{t_2}$, where λ_{t_2} is the shear wavelength in fused quartz (Fig. 11). They are also compared to the losses deduced from the imaginary parts of the two solutions (9). The attenuation due to the leakage into the Plexiglas is higher for the first mode so that the second mode is seen on a longer distance along the interface.

Finally, the amplitudes in the x_2 direction of these modes in each medium are determined from the experimen-

Matteï *et al.*: Experimental investigations of acoustic modes   1537
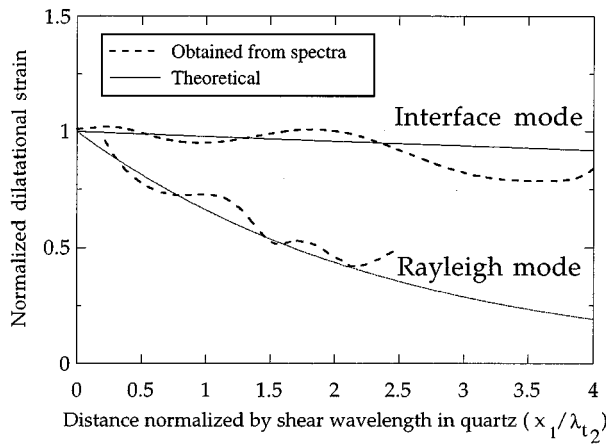
FIG. 11. Respective attenuation along the interface of each mode deduced from the spectra in function of distance normalized by the shear wavelength in fused quartz. The continuous lines are deduced from the imaginary parts of the wave vectors solutions of the characteristic equation.

FIG. 13. Field amplitude deduced from the spectra in the direction perpendicular to the interface for the interface mode as a function normalized by the shear wavelength in fused quartz. The continuous lines are deduced from the wave vector solution of the characteristic equation.

tal data. They are represented in Fig. 12 for the first mode and in Fig. 13 for the second mode. The experimental data are compared to the theoretical inhomogeneities deduced from the calculation of the dilatation strain field. Both modes are confined inside fused quartz and the physical difference appears when considering the field inhomogeneities inside Plexiglas. The leakage of the first mode's longitudinal potential (proportional to the strain dilatation) is clearly seen: The amplitude in Plexiglas is higher with the distance to the interface. This mode corresponds with the solution called Rayleigh wave by Pilant. On the other hand, the second mode's longitudinal potential is slightly confined at the interface. This mode is analog to the solution called Interface wave by Pilant. The interface wave reradiates only in Plexiglas through the shear part of the potentials. Finally the dilatation strain measurements permit identification without ambiguity of the two solutions described by Pilant, the Rayleigh wave and the Interface wave, simultaneously detected for these material combinations.

IV. CONCLUSION

Direct measurements of several interface waves propagating along a solid–solid interface by optical means have been achieved. The detection technique, already used to detect Rayleigh and Lamb waves, is based on the interferometric detection of the optical phase shift of a laser beam induced by the crossing of an acoustical dilatation strain field inside a transparent medium. The classic case of a Stoneley wave propagating without loss along an interface has been illustrated. Velocity and confinement at the interface are determined from the scans in both normal and parallel directions to the interface. When the combination of materials does not support pure real celerity modes, one or two modes propagates with loss along the interface. These two modes exist simultaneously in the case of the combination fused quartz–Plexiglas. This situation has been experimentally illustrated. The features of the interferometric detection method have allowed a complete and straightforward characterization of these two modes. It gives access to more than just the velocities of the waves but also to their field inhomogeneities. These experimental results show the ability of this new optical detection method to characterize interface wave propagation because it leads to direct and quantitative measurements of the dilatational strain field inside transparent materials.

ACKNOWLEDGMENT

FIG. 12. Field amplitude deduced from the spectra in the direction perpendicular to the interface for the Rayleigh mode as a function normalized by the shear wavelength in fused quartz. The continuous lines are deduced from the wave vector solution of the characteristic equation.

[1] *Research Techniques in Nondestructive Testing*, edited by R. S. Sharpe (Academic, New York, 1985).
[2] *Elastic Waves in the Earth*, edited by W. L. Pilant (Elsevier, New York, 1979).
[3] L. J. Pirak-Nolte, L. R. Myer, and N. G. Cook, "Transmission of seimic waves across single natural fractures," J. Geophys. Res. **95**, 8617–8638 (1990).
[4] S. I. Rokhlin and Y. J. Wang, "Analysis of boundary conditions for elastic wave interaction with an interface between two solids," J. Acoust. Soc. Am. **89**, 503–515 (1991).
[5] R. Stoneley, "Elastic waves at the surface of separation of two solids," Proc. R. Soc. London, Ser. A **106**, 416–428 (1924).

[6] J. G. Sholte, ''The range of existence of Rayleigh and Stoneley waves,'' Mon. Not. R. Astron. Soc., Geophys. Suppl. **5**, 120–126 (1947).

[7] G. S. Murty, ''A theoretical model for the attenuation and dispersion of Stoneley waves at the loosely interface of elastic half-spaces,'' Phys. Earth. Plan. Int. **11**, 65–79 (1975).

[8] D. A. Lee and D. M. Corbly, ''Use of interface waves for nondestructive inspection,'' IEEE Trans. Sonics Ultrason. **SU-24** (3), 206–212 (1977).

[9] R. O. Claus and C. H. Palmer, ''Direct measurement of ultrasonic Stoneley wave,'' Appl. Phys. Lett. **31**, 547–548 (1977).

[10] C. B. Scruby and L. E. Drain, *Laser Ultrasonics* (Adam Hilger, New York, 1990).

[11] X. Jia, Ch. Matteï, and G. Quentin, ''Analysis of optical interferometric measurements of guided acoustic waves in transparent solid,'' J. Appl. Phys. **77**, 5528–5538 (1995).

[12] W. L. Pilant, ''Complex roots of the Stoneley-wave equation,'' Bull. Seismol. Soc. Am. **62**, 285–299 (1972).

[13] Ch. Matteï, X. Jia, and G. Quentin, ''Measurement of Rayleigh wave strains inside a transparent solid by optical interferometry,'' Acustica united with Acta Acustica **2**, 65–67 (1994).

[14] A. Korpel, *Acousto-optics* (Marcel Dekker, New York, 1988).

[15] A. Yariv and P. Yeh, *Optical Waves in Crystals* (Wiley, New York, 1984).

[16] D. Royer and E. Dieulesaint, ''Optical detection of sub-ångström transient mechanical displacements,'' Proc. IEEE Ultrason. Symp. 527–529 (1986).

[17] B. Poirée and F. Luppé, ''Evanescent plane waves and the Scholte-Stoneley interface wave,'' J. Acoust. **4**, 575–588 (1991).

New methods to characterize radially polarized piezoelectric ceramic cylindrical shells of finite length

D. D. Ebenezer and A. J. Sujatha

Transducer Group, Naval Physical and Oceanographic Laboratory, Naval Base, Kochi 682 004, India

New methods to experimentally determine the electroelastic coefficients of radially polarized piezoelectric ceramic cylindrical shells (RPPCCSs) of arbitrary length to radius ratio are presented. The real parts of ϵ_{33}^T, s_{11}^E, s_{12}^E, and g_{31} are explicitly expressed in terms of quantities which can be easily measured. The quantities can be obtained from the measured frequency-dependent input electrical admittances (IEAs) of RPPCCSs with free ends. The theoretical IEA based on the two-dimensional equations of state and the membrane approximation is used to derive the explicit expressions. Some methods cannot be used when the resonance frequencies, which are required as inputs, cannot be identified. This occurs when there are several resonances in a narrow frequency band or losses are high. However, the other methods can be used. The methods are validated by using simulated data. The required input data are computed using the theoretical expression for the IEA and an assumed set of complex electroelastic coefficients. It is then shown that the real parts of the coefficients of RPPCCSs with a small or large length to radius ratio and low or high mechanical losses can be accurately determined by using the input data. The effect of measurement errors on the overall accuracy of the experimentally determined values is also discussed. The methods are further validated by the good agreement over a wide frequency range between measured IEAs and the IEAs computed using the experimentally determined values of coefficients. © *1997 Acoustical Society of America.* [S0001-4966(97)01509-9]

PACS numbers: 43.38.Ar, 43.38.Fx, 43.30.Yj, 43.20.Ye [SLE]

INTRODUCTION

Precise control of the magnitude and phase of the acoustic sensitivity of individual hydrophones deployed in an array is essential for accurate estimation of the direction of arrival of signals. Therefore, accurate methods to characterize the piezoelectric component of the hydrophone are critical.

Radially polarized piezoelectric ceramic cylindrical shells (RPPCCSs) are probably the most widely used active components in hydrophones. Therefore, methods to characterize RPPCCSs are of interest.

The ANSI/IEEE standard[1] can be used to determine the real parts of some of the electroelastic coefficients of piezoelectric bars, thin plates, and disks. Several methods which are not part of the standard have also been proposed to experimentally determine the electroelastic coefficients of a few piezoelectric ceramic components. Holland and EerNisse[2] presented a gain-bandwidth method to determine the real parts of some of the electroelastic coefficients of piezoelectric bars. Later, Holland[3] presented a method to also determine the imaginary parts. Smits[4] used an iterative method to determine the complex electroelastic coefficients of bars. Martin[5] presented a method to characterize a stack of longitudinally polarized piezoelectric ceramic cylinders by using a one-dimensional model of the stack. A method to characterize bars, plates, and disks was discussed[6,7] quite recently.

Several methods to characterize RPPCCSs have been proposed. Onoe and Jumonji[8] used approximate equations to characterize RPPCCSs with either infinite or very small length to radius ratio. Ricketts[9] presented an impulse method to determine the complex electroelastic coefficients of RPPCCSs with very small length to radius ratios. Wang[10] has indicated that the electroelastic coefficients of RPPCCSs with arbitrary length to radius ratio can be experimentally determined by solving certain simultaneous equations. Zalesak and Rogers[11] proposed a method to characterize RPPCCSs of arbitrary length to radius ratio. The free-field voltage sensitivities at high and low frequencies are used as inputs in their method.

In this paper, new methods to determine the electroelastic coefficients of thin radially polarized piezoelectric ceramic cylinder shells of arbitrary length to radius ratio are presented. The real parts of the coefficients are explicitly expressed in terms of quantities which can be easily measured. The quantities can be obtained from the measured frequency-dependent IEAs of RPPCCSs with stress-free ends. The theoretical IEA based on the two-dimensional equations of state and the membrane approximation are used to derive the explicit expressions.

ϵ_{33}^T is first expressed in terms of the capacitance at very low frequencies. Then methods I and II to experimentally determine s_{11}^E and s_{12}^E are presented. They are explicitly expressed in terms of resonance frequencies at which bending stiffness is not significant. Next, it is shown that g_{31} can be determined by using methods A, B, and C. Methods A and B are analogous to the standard[1] methods to determine g_{31} of a bar. The anti-resonance frequencies and the slope of the input electrical reactance, respectively, are used in these methods. The input electrical susceptances at frequencies below

the lowest resonance frequency are used to determine g_{31} in method C.

Some methods, which require resonance frequencies as inputs, cannot be used when resonances which are close to each other in frequency space merge, making it impossible to identify all of them. This occurs when the length to radius ratio is small or the imaginary part of the electroelastic coefficients are large. However, the other methods can be used.

The methods are validated by using simulated data. Data computed using the theoretical expression for the IEA and an assumed set of complex electroelastic coefficients are used as inputs. It is then shown that the real parts of the four coefficients of RPPCCSs with a small or large length to radius ratio and low or high mechanical losses obtained using the input data are in good agreement with the assumed set. The effect of measurement errors on the overall accuracy of the experimentally determined values is also discussed. The methods are further validated by using experimental data. It is shown that the IEAs of RPPCCSs computed using the experimentally determined values of coefficients are in good agreement with the measured IEAs in a wide frequency range.

I. CHARACTERIZATION

A. Introduction

In principle, frequency-dependent electroelastic coefficients can be determined by using measured quantities such as IEA, acceleration sensitivity, mechanical impedance, etc. at each frequency and using the appropriate equations to determine the electroelastic coefficients. For example, it is known that ϵ_{33}^T of a RPPCCS can be determined by using the measured value of IEA at low frequencies. Similarly, g_{31} of a RPPCCS can be determined by using the measured value of acceleration sensitivity[12] at low frequencies. The approach can, in principle, be used to obtain values with very high accuracy.

Our primary interest is in characterizing a large number of RPPCCSs and identifying those with unacceptable values of g_{31} or ϵ_{33}^T. The approach indicated in the previous paragraph and the one in Ref. 11 are not suitable for this purpose because a long time will be required to characterize each cylinder. Further, it is sufficient to assume that the coefficients are independent of frequency and obtain accuracies of a few percent since variations in the sensitivity of the hydrophone caused by other factors are generally much greater. Hence a new, easy, and quick approach to characterize RPPCCSs is presented.

Consider a RPPCCS of length L, mean radius a, and thickness h, as shown in Fig. 1. The outer and inner curved surfaces are electroded. The thickness is much less than the radius and the length. It can be shown by using the membrane approximation that the IEA, $Y(\omega)$, under stress-free boundary conditions can be expressed[13,14] as
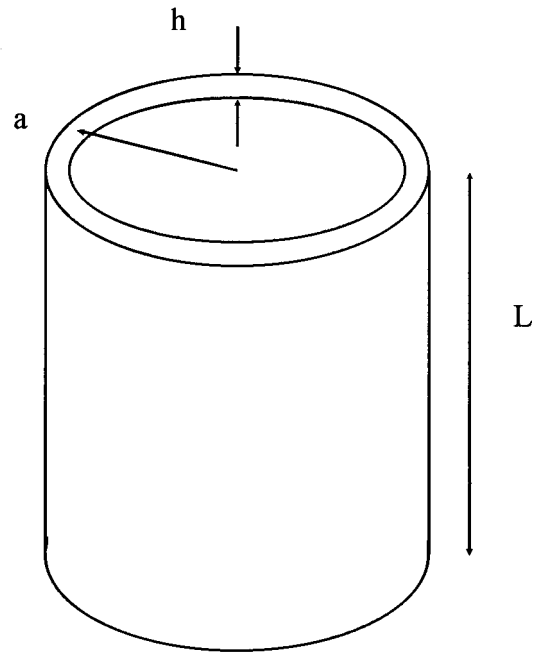
FIG. 1. A thin radially polarized piezoelectric ceramic cylindrical shell. The inner and outer curved surfaces are electroded.

$$Y(\omega) = G(\omega) + jB(\omega)$$

$$= j\omega C_1 \left[\frac{\alpha_4}{\alpha_1} + k_{31}^2 \frac{\alpha_3^2 \tan(\mathbb{K}L/2)}{\alpha_1 \alpha_2 (\mathbb{K}L/2)} \right], \quad (1a)$$

where $G(\omega)$ and $B(\omega)$ are the input electrical conductance and susceptance, respectively, ω is the angular frequency,

$$C_1 = 2\pi a L \epsilon_{33}^T / h, \quad (1b)$$

$$\mathbb{K} = \omega / c_R, \quad (1c)$$

$$c_R^2 = c^2 \alpha_2 / \alpha_1, \quad (1d)$$

$$c^2 = (\rho s_{11}^E)^{-1}, \quad (1e)$$

$$\alpha_1 = 1 - (1 - \sigma^2)\Omega^2, \quad (1f)$$

$$\alpha_2 = 1 - \Omega^2, \quad (1g)$$

$$\Omega = \omega a / c, \quad (1h)$$

$$\sigma = -s_{12}^E / s_{11}^E, \quad (1i)$$

$$\alpha_3 = 1 - (1 + \sigma)\Omega^2, \quad (1j)$$

$$\alpha_4 = 1 - k_{31}^2 - (1 + \sigma)(1 - \sigma - 2k_{31}^2)\Omega^2, \quad (1k)$$

$$k_{31}^2 = d_{31}^2 (\epsilon_{33}^T s_{11}^E)^{-1}, \quad (1l)$$

and ρ is the density of the RPPCCS. The IEA can be easily and accurately measured. Therefore, explicit expressions for the electroelastic coefficients are derived here in terms of this quantity after assuming that the coefficients in Eqs. (1) are real and independent of frequency. [The IEA of a lossy cylinder *can* be determined by using complex coefficients and Eqs. (1) above.]

B. Determination of ϵ_{33}^T

At low frequencies, i.e., when Ω and $\mathbb{K}L \ll 1$, Eq. (1a) reduces to

$$Y(\omega) \cong j\omega 2\pi a L \epsilon_{33}^T/h. \tag{2}$$

It follows from Eq. (2) that the value of ϵ_{33}^T, which is assumed to be real, can be obtained by using the measured value of $B(\omega)$ at a low frequency and

$$\epsilon_{33}^T = C_{lf} h/(2\pi a L), \tag{3a}$$

where

$$C_{lf} = B(\omega)/\omega. \tag{3b}$$

C. Determination of s_{11}^E and s_{12}^E

When the coefficients are real, it can be seen from Eq. (1a) that the absolute value of $Y(\omega)$, $|Y(\omega)|$, is infinity when

$$\alpha_2 = 1 - \Omega^2 = 0, \tag{4a}$$

as well as when

$$\mathbb{K}L/2 = n\pi/2 \quad n = 1,3,5,\dots. \tag{4b}$$

However, $|Y(\omega)|$ is not infinity when $\alpha_1 = 0$. It can be shown that the resonance frequencies at which Eqs. (4a) and (4b) are satisfied can be expressed as

$$f_1^r = (\rho s_{11}^E (2\pi a)^2)^{-0.5} \tag{5a}$$

and

$$f_1^{ni} = f_1^r \{[1 + \beta \mp ((1-\beta)^2 + 4\sigma^2\beta)^{0.5}]/$$
$$(2(1-\sigma^2))\}^{0.5} \quad i = 1,2, \tag{5b}$$

respectively, where

$$\beta = (n\pi a/L)^2. \tag{5c}$$

In Eq. (5b) f_1^{ni} is determined by using the minus and plus signs when $i = 1$ and 2, respectively. The subscript 1 indicates that the admittance is maximum at these frequencies. It is noted that f_1^r is sometimes referred to as the hoop resonance frequency and that the resonance frequencies f_1^{ni}, $i = 1,2$, are referred to as being in the lower and upper branches, respectively.

In an ideal lossless cylinder, the electroelastic coefficients are real. It therefore follows that the resonance frequencies f_1^r and f_1^{ni} are real. In a lossy cylinder, the coefficients are complex, $|Y(\omega)|$ is always finite, and three frequencies[15] are associated with each resonance frequency of a lossless cylinder. These are the frequencies at which $|Y(\omega)|$ and G are maximum and B is zero. Let the frequencies at which G reaches a maximum be f_s^r and f_s^{ni}. The real parts of s_{11}^E and s_{12}^E can be determined by substituting[15] the measured values of f_s^r and f_s^{ni} for f_1^r and f_1^{ni}, respectively, in Eqs. (5) because they are approximately equal.

1. Method I

In this method, the measured value of f_s^r and the measured value of one of the resonance frequencies f_s^{ni} are used. It can be shown after rearranging Eqs. (5a) and (5b) that

$$s_{11}^E = [\rho(2\pi f_s^r a)^2]^{-1} \tag{6a}$$

and

$$\sigma^2 = 1 - \psi(1+\beta) + \psi^2\beta, \tag{6b}$$

where

$$\psi = \left(\frac{f_s^r}{f_s^{ni}}\right)^2, \tag{6c}$$

and f_s^r and f_s^{ni} have been substituted for f_1^r and f_1^{ni}, respectively.

The value of s_{11}^E can be obtained by using Eq. (6a). The value of s_{12}^E can be obtained by using the measured values of f_s^r, f_s^{11}, and Eqs. (6a) and (6b). It can also be obtained by using the measured value of f_s^{31} or some other resonance frequency instead of f_s^{11}. However, since the membrane approximation has been used to develop an expression for the input electrical admittance, the error in the expression for f_s^{n1} will increase with n. It is therefore better, whenever possible, to use the measured value of f_s^{11} to determine σ. If f_s^{11} cannot be clearly identified, then it is highly unlikely that f_s^{n1}, $n > 1$, can be identified. In such cases f_s^{12} may be used.

2. Method II

In this method, the measured values of two resonance frequencies are used. f_s^r is not used. This method is useful when f_s^r cannot be identified because it is very close to other resonance frequencies such as f_s^{n1}, $n > 1$.

After substituting f_s^r and f_s^{ni} for f_1^r and f_1^{ni}, respectively, Eqs. (5) can be rearranged to obtain

$$s_{11}^E = [1 + \beta \mp ((1-\beta)^2 + 4\sigma^2\beta)^{0.5}]/$$
$$[2\rho(2\pi a f_s^{ni})^2(1-\sigma^2)] \tag{7}$$

and

$$\sigma^2 = [\mu^2(1+\beta)^2 - (1-\beta)^2]/(4\beta), \tag{8a}$$

where

$$\mu = \frac{(f_s^{n2})^2 - (f_s^{n1})^2}{(f_s^{n2})^2 + (f_s^{n1})^2}. \tag{8b}$$

The measured values of f_s^{n1} and f_s^{n2} and Eqs. (7) and (8) can be used to determine s_{11}^E and s_{12}^E. Since the membrane approximation has been used, f_s^{11} and f_s^{12} should be used for maximum accuracy.

D. Determination of g_{31}

Three methods, A, B, and C, are now presented to determine g_{31}. After first finding k_{31}, g_{31} can be determined by using

$$g_{31} = k_{31}(s_{11}^E/\epsilon_{33}^T)^{0.5}. \tag{9}$$

Methods A and B are analogous to the standard[5] methods to determine the coupling coefficient of a bar.

1. Method A

It can be shown after rearranging Eq. (1a), and using $Y(\omega) = B(\omega)$ when the coefficients are real, that

$$k_{31}^2 = \left(1 - \frac{B(\omega)}{j\omega C_1}\right)\alpha_1\left(1 - 2(1+\sigma)\Omega^2\right.$$
$$\left. - \frac{\alpha_3^2 \tan(\mathbb{K}L/2)}{\alpha_2(\mathbb{K}L/2)}\right)^{-1}. \tag{10}$$

Then, k_{31}, which is assumed to be real, can be easily determined by using Eq. (10) and the measured value of a frequency at which $B(\omega)$ is zero.

In a lossless cylinder, the absolute value of the input electrical impedance, $|Z(\omega)|$, becomes infinity at the resonance frequencies f_2^r and f_2^{ni}. The subscript 2 indicates that the impedance is maximum at these frequencies. $B(\omega)$ becomes zero at each of these frequencies.

In a lossy cylinder, $|Z(\omega)|$ is always finite and three frequencies[15] are associated with each resonance frequency of a lossless cylinder. These are the frequencies at which $|Z(\omega)|$ and $R(\omega) = \mathrm{Re}[Z(\omega)]$ are maximum and $B(\omega)$ is zero. Let the frequencies at which $R(\omega)$ reaches a maximum be f_p^r and f_p^{ni} and the frequencies at which $B(\omega)$ become zero be f_a^r and f_a^{ni}, respectively. The measured values of f_p^r or f_p^{ni} and Eq. (10) can be used to determine k_{31} since they are approximately equal to f_2^r and f_2^{ni}, respectively. The values of f_a^r or f_a^{ni} can also be used. However, when there are losses, f_a^r or f_a^{ni} may not exist. Even when they exist, f_p^r and f_p^{ni} are closer to f_2^r and f_2^{ni}, respectively. Therefore, when this method is used, it is better to use f_p^r or f_p^{ni}. Further, it is better to use f_p^{1i} than to use f_p^{ni}, $n > 1$. The reasons for this choice are similar to those presented in method I to determine σ. This method, however, may not be useful if there are several resonances in a small frequency band.

2. Method B

In this method, the slope of the input electrical reactance, $dX(\omega)/d\omega$, at the resonance frequency f_s^{ni}, is used to determine k_{31}. After using

$$\tan(\mathbb{K}L/2)\big|_{f=f_1^{ni}} = \pm\infty, \quad n = 1,3,\ldots, \ i = 1,2 \tag{11a}$$

and

$$\frac{dZ(\omega)}{d\omega} = -\frac{1}{Y^2(\omega)}\frac{dY(\omega)}{d\omega} \tag{11b}$$

and

$$\frac{d\mathbb{K}(\omega)}{d\omega} = \frac{1}{c}\left(\frac{\alpha_1}{\alpha_2}\right)^{0.5}\left[1 + \Omega^2\left(\frac{1}{\alpha_2} - \frac{1-\sigma^2}{\alpha_1}\right)\right] \tag{11c}$$

to obtain an expression for $dX(\omega)/d\omega\big|_{f=f_1^{ni}}$, it can be shown that

$$k_{31}^2 = \frac{h}{\epsilon_{33}^T 2\pi a L}\left(\frac{\alpha_1 L}{\alpha_3 2c}\right)^2\left[1 + \Omega^2\left(\frac{1}{\alpha_2} - \frac{1-\sigma^2}{\alpha_1}\right)\right]$$
$$\times\left(\frac{dX(\omega)}{d\omega}\right)^{-1}\bigg|_{f=f_1^{ni}}. \tag{11d}$$

The coupling coefficient can be determined by substituting the measured value of the slope of the reactance at f_s^{ni} in Eq. (11c). It is better, as in methods I and II, to use f_s^{1i}.

3. Method C

In this method, k_{31} and g_{31} are determined by using the previously determined values ϵ_{33}^T, s_{11}^E, and s_{12}^E, Eq. (10), and the measured value of $B(\omega)$ at some frequency. The frequency can be selected after observing that the error in the computed value of g_{31} is equal to the sum of the error due to measurement errors and the error caused by assuming that the coefficients are real while deriving expressions for them.

First, consider the measurement error. After defining

$$\delta B(\omega) = \Delta B(\omega)/B(\omega), \tag{12a}$$

where $\Delta B(\omega)$ is the measurement error and $B(\omega)$ is the correct value, it can be shown by using Eqs. (9) and (10) that

$$\delta g_{31} = \Delta g_{31}/g_{31} = [1 + (1+\eta)\delta B(\omega)/\eta]^{0.5} - 1, \tag{12b}$$

where Δg_{31} is the error in the computed value of g_{31}, and

$$\eta = B(\omega)/(\omega C_1) - 1,$$
$$\cong 0 \quad \omega \to 0. \tag{12c}$$

It can easily be seen from Eqs. (12b) and (12c) that when η is small, i.e., at low frequencies, the error in g_{31} is large even if the measurement error is small.

Second, at frequencies which are very close to the resonance frequencies, the imaginary parts of the coefficients, which are assumed to be zero while deriving expressions for the coefficients, have a significant effect on the measured value of $B(\omega)$. Therefore, the error in the value of g_{31} determined by using the value of $B(\omega)$ at such frequencies is again likely to be large.

It is clear from the above that the optimum frequency is a function of the measurement error as well as the imaginary parts of the coefficients. Some recommendations for the selection of the optimum frequency based on numerical results are made in the next section.

II. RESULTS AND DISCUSSION

In this section, numerical results are first presented to illustrate the theoretical IEAs of RPPCCSs computed using membrane theory. Next, tables of electroelastic coefficients obtained using the IEAs and various methods are discussed. Then, the effect of errors in measurement on the computed coefficients is presented. Finally, experimental data are used to compute the coefficients. The input electrical susceptances computed using these coefficients are then compared with the measured values over a wide frequency range.

The theoretical IEAs of RPPCCSs computed using the complex coefficients in sets I, II, and III of Table I are shown in Figs. 2 and 3. $a = 0.01$ m, $h/a = 0.1$, and $\rho = 7500$ kg/m^3 in the figures. $L/a = 0.5$ and 4 in Figs. 2 and 3, respectively. The real parts of the coefficients are the same in all the sets and the imaginary parts satisfy the conditions in the Appendix. The absolute values of $s_{11}^{E''}$ and $s_{12}^{E''}$ in sets II and III are 10 times those in set I. $s_{11}^{E''}$ and $s_{12}^{E''}$ have the same sign in sets I and II and are of opposite sign in set III.

Some of the resonance frequencies in Figs. 2 and 3 cannot be easily identified. Most of the discussion in the following paragraphs is regarding these frequencies.

TABLE I. Values of electroelastic coefficients used to obtain simulated data.

	Set I	Set II	Set III
$\epsilon_{33}^T/\epsilon_0$	$1300(1-j0.004)$	$1300(1-j0.004)$	$1300(1-j0.004)$
$s_{11}^E(10^{-12}\ \text{m}^2/\text{N})$	$12-j0.06$	$12-j0.6$	$12-j0.6$
$s_{12}^E(10^{-12}\ \text{m}^2/\text{N})$	$-4-j0.06$	$-4-j0.6$	$-4+j0.6$
$g_{31}(10^{-3}\ \text{Vm/N})$	$-10+j0.1$	$-10+j0.4$	$-10-j0.04$

Figure 2(a) can be better appreciated if the resonance frequencies of a corresponding lossless cylinder whose coefficients are the real parts of the coefficients in Table I are known. It can be easily shown that (see Sec. I C) $f_1^r = 53.052$ kHz, $f_1^{11} = 52.975$ kHz, and $f_1^{31} = 53.043$ kHz. It is noted that $f_1^r \cong f_1^{n1}$. It can also be shown that (see Sec. I D) $f_2^r = 55.661$ kHz, $f_2^{12} = 53.032$ kHz, and $f_2^{32} = 53.047$ kHz.

The input electrical conductances computed using sets I, II, and III are shown by solid, dashed, and solid-x lines, respectively, in Fig. 2(a). Each plot has only one peak at 52.983, 52.947, and 53.008 kHz, respectively. After comparing these frequencies with the resonance frequencies of the lossless cylinder, they are called f_s^{11}. f_s^r cannot be identified from Fig. 2(a) because all the resonances in the lower branch have merged with the hoop resonance. It can be seen that there is not much difference between the IEAs computed using sets II and III and shown in Fig. 2(a) and (b).

In Fig. 2(c) and (d), the real and imaginary parts of the IEAs are shown in the neighborhood of the lowest upper branch resonance frequency. In this case there is considerable difference between the IEAs obtained using sets II and III. The frequency at which the peak occurs in each plot of Fig. 2(c) is called f_s^{12}. They are approximately equal to the
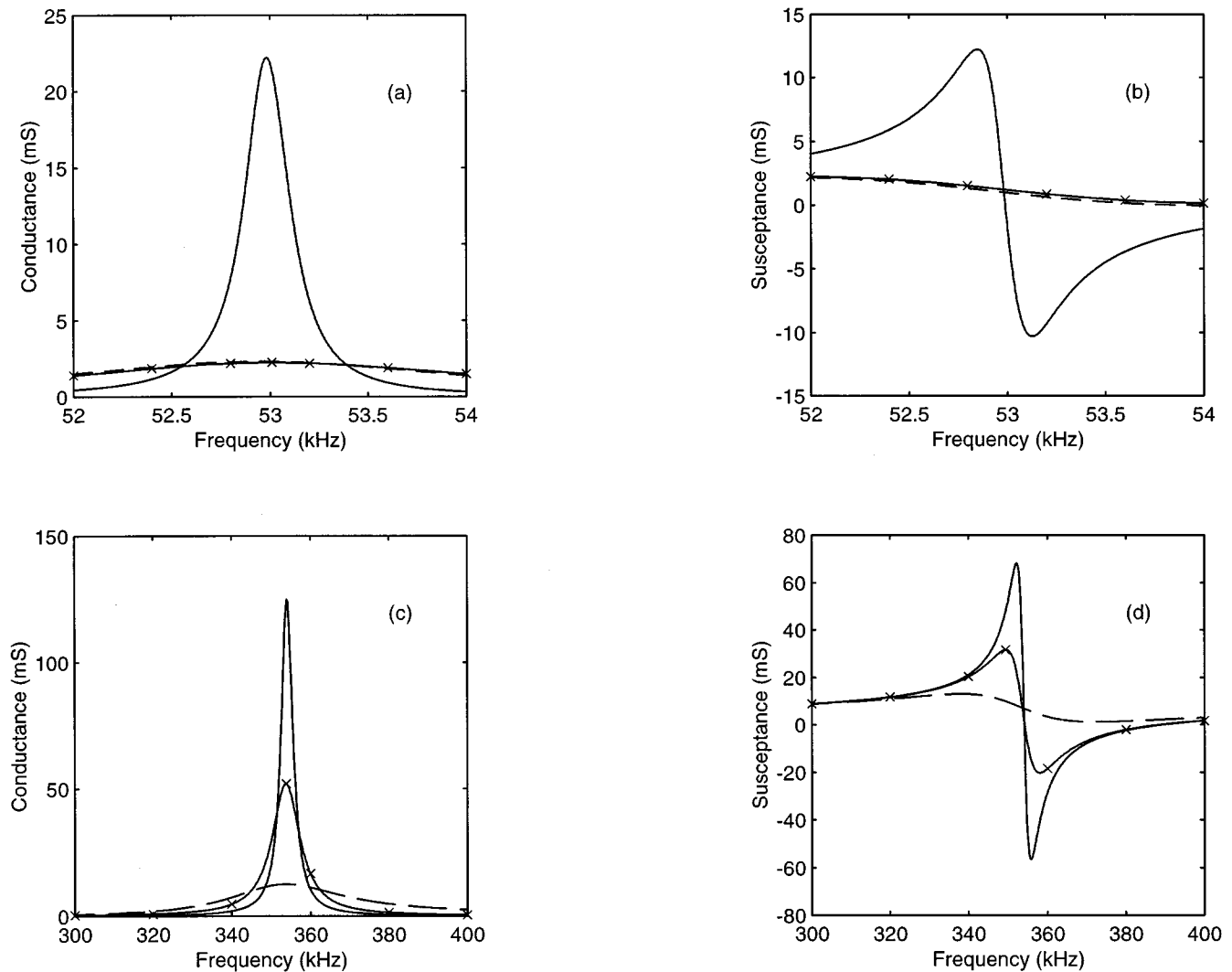
FIG. 2. The theoretical input electrical admittances of radially polarized piezoelectric ceramic cylindrical shells of length 0.005 m, radius 0.01 m, and wall thickness 0.001 m computed using the coefficients in Table I. — set I; --- set II; -×-×-×-× set III. (a) and (c) conductance; (b) and (d) susceptance.
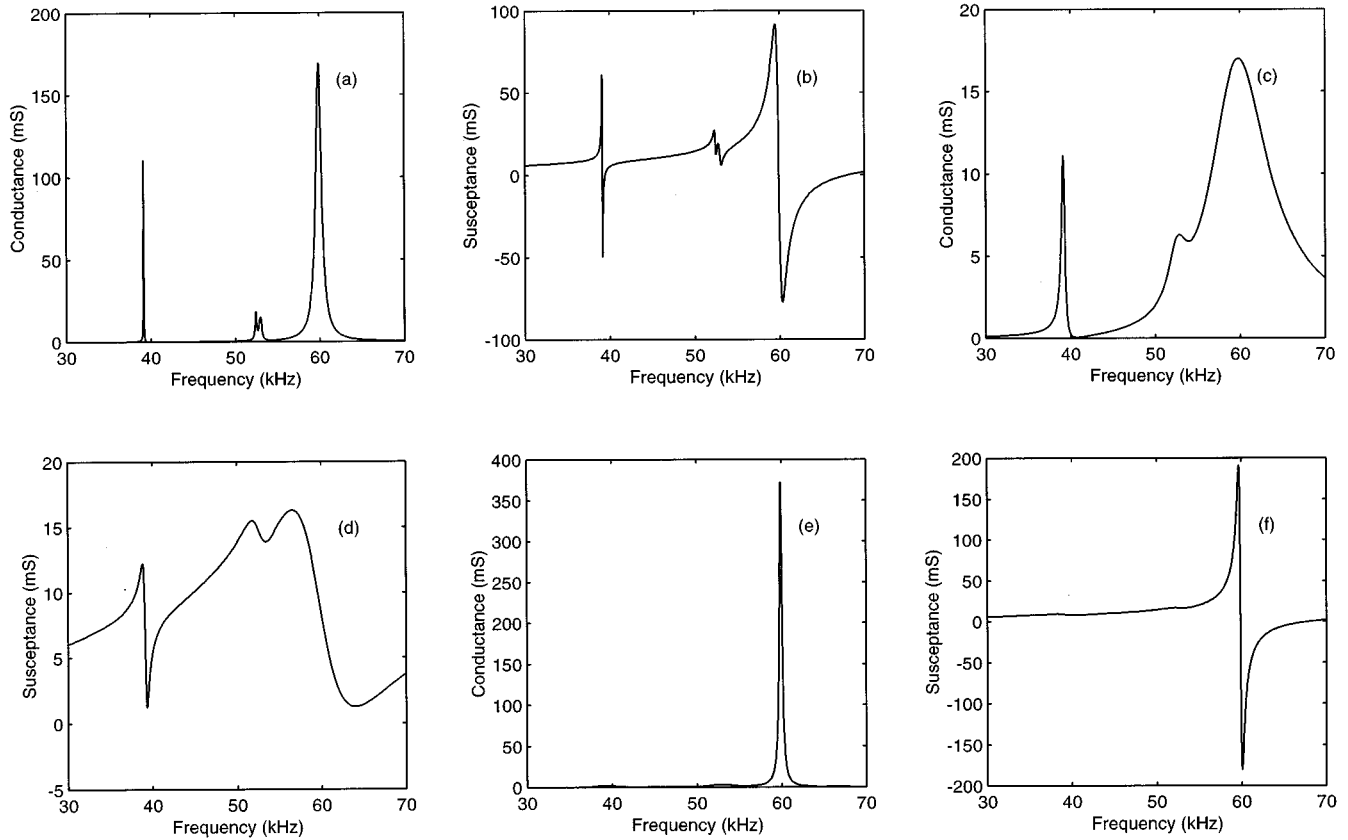
FIG. 3. The theoretical input electrical admittances of radially polarized piezoelectric ceramic cylindrical shells of length 0.04 m, radius 0.01 m, and wall thickness 0.001 m computed using the coefficients in Table I. (a) and (b) set I; (c) and (d) set II; (e) and (f) set III.

frequency $f_1^{12} = 354$ kHz of the corresponding lossless cylinder.

The IEAs of cylinders with $L/a = 4$ are shown in Fig. 3(a)–(f). In this case, the resonance frequencies of the corresponding lossless cylinder are $f_1^r = 53.052$ kHz, $f_1^{11} = 39.148$ kHz, $f_1^{31} = 52.433$ kHz, $f_1^{12} = 59.890$ kHz, $f_2^r = 53.154$ kHz, $f_2^{11} = 39.502$ kHz, and $f_2^{12} = 68.148$ kHz. A frequency resolution which is high enough to ensure that the local maxima and minima are correctly represented has been used to obtain the figures.

When sets I, II, and III are used, peaks occur in the input electrical conductance at 39.147, 39.138, and 39.249 kHz, respectively, as seen in Fig. 3(a), (c), and (e). These frequencies are called f_s^{11}.

When set I is used, the peaks expected at the lower branch resonance frequencies, f_s^{n1} $n > 3$, merge with the expected peak at the hoop mode resonance frequency (f_s^r), as seen in Fig. 3(a). The frequency closest to f_1^r at which a peak occurs, 52.979 kHz, is called f_s^r. When sets II or III are used,

the peaks expected at f_s^{n1} $n > 1$, merge with the expected peak at f_s^r and there is only one peak in the vicinity of f_1^r. This can be seen in Fig. 3(c) and (e). The frequencies, 52.880 kHz and 52.832 kHz, respectively, at which this peak occurs are called f_s^r. The frequencies f_s^{12} can be easily identified in the figure and are equal to 59.886 kHz, 59.838 kHz, and 59.866 kHz when sets I, II, and III, respectively, are used.

The input electrical susceptances are shown in Fig. 3(b), (d), and (f). There are only a few zero crossings, indicating that only some of the frequencies f_a^r and f_a^{ni} exist.

Next, details of an approach which was used to determine the relative merits of the methods to characterize RPPCCSs are presented. The inputs required to compute the coefficients using the various methods are usually obtained experimentally. Here, the IEAs computed using Table I and shown in Figs. 2 and 3 were used. Since there are no measurement errors the differences between the real parts of the electroelastic coefficients used to compute the IEAs and the computed values of the real parts of the coefficients give an

TABLE II. Values of s_{11}^E and s_{12}^E computed using method I. The values of f_s^{11} and f_s^r were used as inputs.

L/a	1/2			4		
Set	I	II	III	I	II	III
$s_{11}^E (10^{-12}$ m²/N$)$	⋯	⋯	⋯	12.033	12.078	12.100
$s_{12}^E (10^{-12}$ m²/N$)$	⋯	⋯	⋯	−3.953	−3.896	−3.740

TABLE III. Values of s_{11}^E and s_{12}^E computed using method II. The values of f_s^{11} and f_s^{12} were used as inputs.

L/a	1/2			4		
Set	I	II	III	I	II	III
$s_{11}^E (10^{-12}$ m²/N$)$	11.997	12.013	11.986	12.001	12.010	11.960
$s_{12}^E (10^{-12}$ m²/N$)$	−3.993	−3.992	−3.950	−4.000	−3.995	−3.948

L/a	1/2			4		
Set	I	II	III	I	II	III
Using f_a^r	−9.98	−8.13	...	...	...	...
f_p^r	−9.99	−9.94	−9.91	−9.92	−16.82	imag
f_a^{11}	...	...	...	−10.01	...	...
f_p^{11}	...	...	...	−10.01	−10.03	−8.64
f_a^{12}	−9.99	...	−9.99	−9.99	...	−10.01
f_p^{12}	−10.00	−9.85	−10.04	−10.00	−9.79	−10.01

TABLE IV. Values of g_{31} in 10^{-3} Vm/N computed using method A and the values of s_{11}^E and s_{12}^E computed using method II.

L/a	1/2			1/4		
Set	I	II	III	I	II	III
Using $dX/d\omega\|_{f=f_s^{11}}$	−11.33	−15.40	−16.85	−10.05	−17.40	imag
$dX/d\omega\|_{f=f_s^{12}}$	−10.04	−16.19	−10.24	−9.98	−15.47	imag

TABLE V. Values of g_{31} in 10^{-3} Vm/N computed using method B and the values of s_{11}^E and s_{12}^E computed using method II.

indication of the errors inherent in the various methods. However, since the computed value of ϵ_{33}^T is used to compute the values of s_{11}^E and s_{12}^E, and the computed values of ϵ_{33}^T, s_{11}^E, and s_{12}^E are used to compute the values of g_{31}, the errors in the computed values of the coefficients are cumulative. Holland and EerNisse[2] and Smits[4] have used the same approach to validate their methods of characterization of bars.

The values of s_{11}^E and s_{12}^E computed using method I are shown in Table II. The resonance frequencies f_s^{11} and f_s^r were used. When $L/a=0.5$, f_s^r does not exist. Therefore, method I cannot be used to find the coefficients. It can be seen, however, that the coefficients can be determined with reasonable accuracy when $L/a=4$ for all three sets. Table II shows that if the coefficients of a cylinder with $L/a=4$ are actually those shown in set I of Table I, i.e., s_{11}^E and s_{12}^E are $12-j0.06$ and $-4-j0.06$ (10^{-12} m^2/N), respectively, then the values of the coefficients obtained by using method I will be 12.033 and −3.953 (10^{-12} m^2/N), respectively.

The values of s_{11}^E and s_{12}^E computed using method II are shown in Table III. In contrast to method I, method II can be used to compute the coefficients when $L/a=0.5$ also. When set I is used, the errors in the computed values of s_{11}^E and s_{12}^E are 0.03% and 0.2%, respectively, when $L/a=0.5$ and 0.008% and almost 0% respectively when $L/a=4$. Method II is recommended since it can be used for small as well as large L/a and it is more accurate than method I when set I or II is used.

The values of g_{31} computed using method A are shown in Table IV. Several resonance frequencies do not exist because of losses. However, for the RPPCCSs considered in the table, g_{31} can be computed by using three or more resonance frequencies for each cylinder. In each case, a majority of the computed values of g_{31} are approximately equal to

each other. Thus it is possible to judge which of the computed values are reasonably accurate, without using the knowledge that the correct value of the real part of g_{31} as shown in Table I is -10 (10^{-3} Vm/N). It can also be seen from Table IV that the values obtained using f_p^r, f_p^{11}, and f_p^{12} are as accurate or more accurate than those obtained using f_a^r, f_a^{11}, and f_a^{12}, respectively, except in one case. When $L/a=0.5$ and set III is used, the value of g_{31} obtained using f_a^{12} is more accurate than that obtained using f_p^{12} because of the cumulative errors discussed earlier in this section.

The values of g_{31} computed using method B are shown in Table V. The values are reasonably accurate only when the losses are low. Therefore, method B is not considered to be very useful.

The values of g_{31} computed using method C are shown in Table VI. The values of B at an arbitrary frequency are used in this method. The frequencies at which B was computed using Eq. (1) for each L/a are also shown in the table. They are approximately equal to 0.5, 0.7, 0.9, 0.95, and 0.98 times f_s^{11} for both values of L/a. It can be seen from the table that when the actual coefficients of the cylinder are given by set I, i.e., the imaginary parts of the coefficients are small, and the values of s_{11}^E and s_{12}^E are computed by using method II, the value of g_{31} can be computed with an error of about 0.1% unless the frequency at which B is computed is $0.98f_s^{11}$. Even when the losses are higher, as in sets II and III, g_{31} can be determined with a maximum error of about 0.2% and 3% by using the computed value of B at $0.5f_s^{11}$ and $0.9f_s^{11}$, respectively, for the cases considered. Therefore, method C is recommended for computing g_{31}.

Next consider the effect shown in Table VII of a +1% error in the measured values of B on the values of g_{31} computed using method C. The percentage errors in the values of g_{31} were computed using Eqs. (12a) and (12b). The errors are shown for the same frequencies considered in Table VI. It can be seen from the table that the error is almost indepen-

TABLE VI. Values of g_{31} in 10^{-3} Vm/N computed using method C and the values of s_{11}^E and s_{12}^E computed using method II.

L/a	1/2			1/4			
Set	I	II	III	Set	I	II	III
Freq. (kHz)		g_{31}		Freq. (kHz)		g_{31}	
26.5	−10.01	−10.01	−10.01	19.5	−9.99	−9.98	−9.99
37.0	−10.00	−9.96	−10.00	27.5	−9.99	−9.95	−10.00
47.5	−10.00	−9.68	−9.82	35.0	−10.0	−9.86	−9.92
50.0	−10.00	−9.07	−9.30	37.0	−10.0	−9.72	−9.51
52.0	−9.94	−5.71	−6.16	38.5	−9.98	−8.82	−6.72

D. D. Ebenezer and A. J. Sujatha: Piezoelectric ceramic shells

TABLE VII. Percentage error in the value of g_{31} computed using method C due to a +1% error in the measured value of $B(\omega)$. All values are for RPPCCSs whose coefficients are given in set I.

$L/a = 1/2$		$L/a = 4$	
Freq. (kHz)	Error (%)	Freq. (kHz)	Error (%)
26.5	14.9	19.5	13.9
37.0	5.8	27.5	6.3
47.5	1.8	35.0	2.9
50.0	1.2	37.0	2.1
52.0	0.7	38.5	1.2

dent of L/a but strongly dependent on the frequency. As discussed previously, μ is small at $0.5f_s^{11}$, and so the error in g_{31} is large. Therefore, the measured value of B at a frequency closer to f_s^{11} should be used when the measurement error is large.

Next, a comparison of the experimental IEAs and the IEAs computed using experimentally determined values of the coefficients is presented. The dimensions of two RPPCCSs used in the comparison are shown in Table VIII. The measured values of the input electrical susceptances are shown in Fig. 4(a) and (b) by dashed lines. All measurements were done with a HP 4192A impedance analyzer using a frequency resolution of 5 Hz. It can be inferred by comparing Fig. 4 with Figs. 2 and 3 that the imaginary parts of the coefficients of the two RPPCCSs are not large. After repeating the measurements several times, it was found that the difference between the largest and smallest values of $B(\omega)$ at a particular frequency was of the order of 1%. Therefore, g_{31} was obtained by using the measured value of B at about $0.95f_s^{11}$.

The theoretical input electrical susceptances of the two RPPCCSs calculated using Eq. (1) and the experimentally determined coefficients are also shown in Fig. 4(a) and (b), respectively, by a solid line. The theoretical and measured values have been plotted at the same set of frequencies. The

TABLE VIII. Dimensions of two RPPCCSs whose input electrical susceptances are shown in Fig. 4(a) and (b), respectively.

L (mm)	76.4	76.1
a (mm)	35.7	17.5
t (mm)	5.1	3.3

measured values of f_s^{11}, f_s^r, and f_s^{12} are 14.150 kHz, 14.875 kHz, and 23.935 kHz, respectively for the first cylinder and 21.040 kHz, 30.200 kHz, and 33.215 kHz, respectively, for the second cylinder. Since method II was used to determine s_{11}^E and s_{12}^E, it can be seen from the figures that there is excellent agreement between the theoretical and measured values in the neighbourhood of f_s^{11} and f_s^{12}. The agreement over the entire frequency range—except near f_s^{n1}, $n \geq 3$—is quite good. This indicates that it is reasonable to assume that the coefficients are independent of frequency.

III. CONCLUSIONS

Several new methods to experimentally determine the real parts of electroelastic coefficients have been presented.

The methods were first validated and compared using simulated data. It was shown that s_{11}^E and s_{12}^E can be most accurately determined by using method II. Method I is less accurate and cannot be used when the hoop resonance frequency cannot be identified. This occurs when the length to radius ratio is small or the imaginary parts of the coefficients are large. Three methods were presented to determine g_{31}. The frequencies at which the susceptance becomes zero and the slopes of the reactances at the frequencies at which conductance is maximum are used in methods A and B, respectively. The two methods are similar to the corresponding ANSI methods to characterize bars and are less accurate than method C when two or more resonances occur in a narrow frequency band. The input electrical susceptance at a selected frequency which is leser than the lowest resonance frequency is used in method C. The value of g_{31} determined
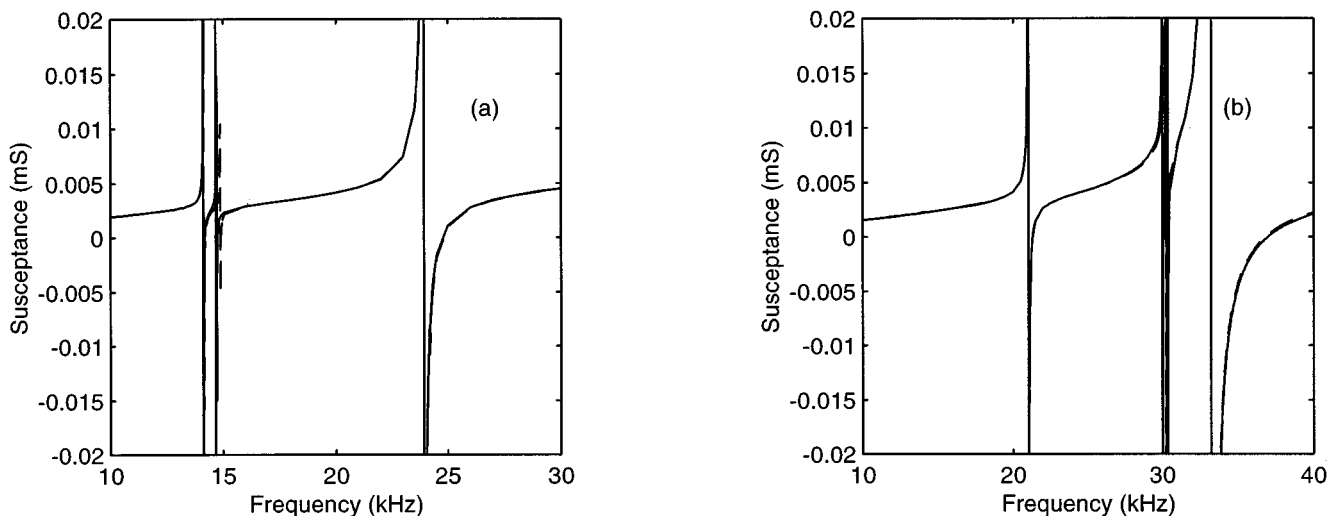
FIG. 4. Input electrical susceptances computed using Eq. (1) and experimentally determined values of coefficients —— and measured values -----. (a) RPPCCS with length=76.4 mm and radius=35.7 mm. (b) RPPCCS with length=76.1 mm and radius=17.5 mm.

experimentally by using this method is quite accurate when the selected frequency is approximately equal to half the lowest resonance frequency and there are no measurement errors. However, when the measurement error is of the order of 1%, and the imaginary parts of the coefficients are small, it is better to select a frequency which is about 0.95 times the lowest resonance frequency. It was shown that the real parts of the coefficients of cylinders with small or large length to radius ratios can be quite accurately determined even when the imaginary parts of the coefficients are large.

The methods were also validated using experimental data. It was shown that the measured input electrical susceptances are in good agreement over a wide frequency band with those computed using the experimentally determined coefficients.

APPENDIX: CONDITIONS SATISFIED BY IMAGINARY PARTS OF COEFFICIENTS

It is well known that electroelastic coefficients are complex. Holland and EerNisse[16] have shown that the negatives of the imaginary parts of ϵ_{33}^T, s_{11}^E, s_{12}^E, and d_{31} which are denoted by $\epsilon_{33}^{T''}$, $s_{11}^{E''}$, $s_{12}^{E''}$, and d_{31}'', respectively, satisfy the following conditions:

$$s_{11}^{E''}, \quad \epsilon_{33}^{T''} > 0, \tag{A1}$$

$$s_{11}^{E''} \geq |s_{12}^{E''}|, \tag{A2}$$

$$s_{11}^{E''} \ \epsilon_{33}^{T''} \geq (d_{31}'')^2, \tag{A3}$$

and

$$\epsilon_{33}^{T''}(s_{11}^{E''} + s_{12}^{E''}) \geq 2(d_{31}'')^2. \tag{A4}$$

[1] T. R. Meeker, ''Publication and proposed revision of ANSI/IEEE standard 176-1987 ANSI/IEEE standard on piezoelectricity,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 717–772 (1996).

[2] R. Holland and E. P. EerNisse, ''Accurate measurements of coefficients in a ferroelectric ceramic,'' IEEE Trans. Sonics Ultrason. **SU-16**, 173–181 (1969).

[3] R. Holland, ''Measurement of piezoelectric phase angles in a ferroelectric ceramic,'' IEEE Trans. Sonics Ultrason. **SU-17**, 123–124 (1970).

[4] J. G. Smits, ''Iterative method for accurate determination of the real and imaginary parts of the material coefficients of piezoelectric ceramics,'' IEEE Trans. Sonics Ultrason. **SU-23**, 393–402 (1976).

[5] G. E. Martin, ''New standard for measurements of certain piezoelectric ceramics,'' J. Acoust. Soc. Am. **35**, 925 (1963).

[6] M. Brissaud, ''Characterization of piezoceramics,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 603–617 (1991).

[7] M. B. Moffett and D. Ricketts, ''Comments on Characterization of piezoceramics,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **38**, 825 (1993); and M. Brissaud, ''Author's Reply,'' *ibid.* **38**, 826–827 (1993).

[8] M. Onoe and H. Jumonji, ''Useful formulas for piezoelectric ceramic resonators and their application to measurement of parameters,'' J. Acoust. Soc. Am. **41**, 974–980 (1967).

[9] D. Ricketts, ''Report on the measurement and evaluation of the dynamic properties of sonar transducer materials. Part II. Mathematical model for the analysis of the dynamic properties of sonar transducer materials using impulse testing techniques,'' Raytheon Co., Submarine Signal Division, RI (15 November 1982).

[10] H. Wang, ''On the tangentially and radially polarized piezoceramic thin cylindrical tube transducers,'' J. Acoust. Soc. Am. **79**, 164–176 (1986).

[11] J. F. Zalesak and P. H. Rogers, ''Method of determining piezoelectric constants of ceramic rings,'' U.S. Patent No. 3,954,015, May 4, 1976.

[12] D. D. Ebenezer, ''Acceleration sensitivity of radially polarized piezoceramic cylinders,'' in *Proceedings of the 3rd International Congress on Air and Structure Borne Sound and Vibration*, edited by M. J. Crocker (International Scientific Publications, Auburn, AL, 1994), Vol. 2, pp. 883–890.

[13] J. F. Haskins and J. L. Walsh, ''Vibrations of ferroelectric cylindrical shells with transverse isotropy. I. Radially polarized case,'' J. Acoust. Soc. Am. **29**, 729–734 (1957).

[14] D. D. Ebenezer, ''Three port parameters and equivalent circuit of radially polarized piezoelectric ceramic cylinders,'' J. Acoust. Soc. Am. **99**, 2908–2912 (1996).

[15] Reference 1, Sec. 6.4.2.

[16] R. Holland and E. P. EerNisse, *Design of Resonant Piezoelectric Devices*, Research Monograph No. 56 (MIT, Cambridge, 1969), pp. 12–16.

Eigenfunction analysis of radially polarized piezoelectric cylindrical shells of finite length

D. D. Ebenezer and Pushpa Abraham

Transducer Group, Naval Physical and Oceanographic Laboratory, Thrikkakara, B.M.C. P.O., Kochi 682 021, India

An eigenfunction approach is presented to determine the response of radially polarized piezoelectric cylindrical shells of finite length to electrical excitation. The equations of motion of the piezoelectric cylinder are first derived by using the membrane approximation. Then they are solved by expressing the displacement distribution as the sum of the static solution and a weighted sum of a complete set of functions—the eigenfunctions of the short-circuited cylinder. The static solution is necessary to exactly satisfy the boundary conditions of the electrically excited cylinder. Moreover, in some cases, the series solution does not converge to the correct displacement distribution when some other terms, instead of the static solution, are used to satisfy the boundary conditions. The weights or displacement coefficients are determined by using the orthogonal property of the eigenfunctions. Finally, the input electrical admittance is determined by using the equation of state and the displacement distribution. Analyses of cylinders with various boundary conditions are presented to illustrate the approach. Numerical results are used to show that the series solutions converge very rapidly to the closed form solutions. © *1997 Acoustical Society of America.* [S0001-4966(97)01609-3]

PACS numbers: 43.38.Ar, 43.38.Fx, 43.40.Ey, 43.20.Ks [SLE]

INTRODUCTION

Closed form expressions for the vibratory responses of elastic cylindrical shells of finite length to nonuniform excitations are, in general, not known. However, series expressions for the responses in air[1] as well as underwater[2,3] have been developed by using eigenfunction methods.

Closed form expressions for the vibratory and electrical responses to electrical excitations of radially polarized piezoelectric cylindrical shells[4,5] have also been developed. This has been possible because, as shown in this paper, the governing equations have exactly the same form as those for elastic cylindrical shells with uniform excitation. Eigenfunction methods are, therefore, not necessary to determine the responses. However, the development of an eigenfunction method to determine the responses may lead to methods to determine the complex coefficients[6] of piezoelectric shells, determine their responses to electrical, acoustic, and mechanical excitations under fluid-loaded conditions and to more nearly accurate methods to analyze other vibrators.

In this article, an eigenfunction approach is presented to analyze electrically excited thin radially polarized piezoelectric cylindrical shells of finite length. In this approach, the equations of axisymmetric motion for the piezoelectric cylinder are first obtained by using the membrane approximation. When the membrane approximation is used, only one boundary condition can be specified at each end of the cylinder. It is assumed that the boundary conditions are independent of frequency. The displacement distribution is then expressed as the sum of the static displacement distribution when a steady or direct voltage is applied and a weighted sum of eigenfunctions. The eigenfunctions are obtained by using the homogeneous equations of motion and the boundary conditions of the short-circuited cylinder. Since the static

displacement distribution satisfies the boundary conditions of the excited cylinder, it follows that the displacement distribution exactly satisfies the boundary condition at all frequencies. The weights or displacement coefficients are determined by using the orthogonal property of the eigenfunctions which form a complete set of functions. In some cases, the eigenfunctions form a normwise complete set of functions and not a pointwise complete set of functions. In such cases, the displacement distribution may not converge to the correct values at the ends of the cylinder even though the boundary conditions are exactly satisfied. This, however, usually does not cause any difficulty since the correct values can be determined by extrapolation. Finally, the input electrical admittance is determined by using the displacement distribution. Analyses of cylinders with various boundary conditions are presented to illustrate the approach and to show that the series solutions converge very rapidly to the closed form solutions.

I. THEORETICAL ANALYSIS

A. Governing equations

Consider a radially polarized piezoelectric cylindrical shell of length L, and mean radius a, as shown in Fig. 1. The wall thickness, h, of the cylinder is assumed to be much smaller than its radius and length. The displacement distribution and the input electrical admittance of the cylinder when it is electrically excited are of interest.

The momentum equations can be expressed in a cylindrical coordinate system as

$$\frac{\partial T_z(z,\omega)}{\partial z} = -\rho\omega^2 U(z,\omega) \tag{1a}$$
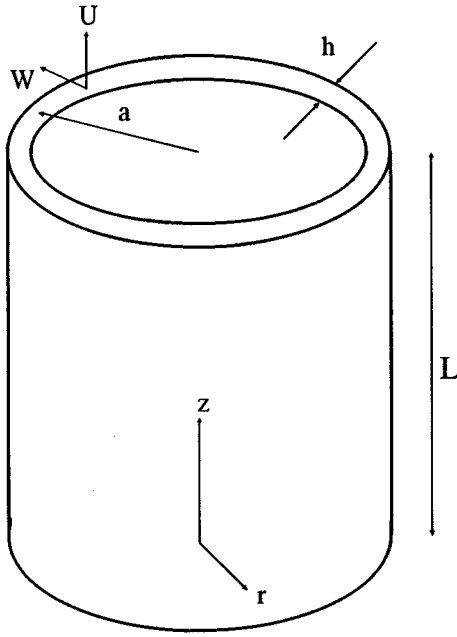
FIG. 1. A thin radially polarized piezoelectric cylindrical shell.

and

$$T_\theta(z,\omega) = \rho\omega^2 a W(z,\omega), \tag{1b}$$

where ω is the angular frequency. U and W are the axial and radial components of displacement, and T_z and T_θ are normal stresses which act in the axial and tangential directions, respectively. Positive values of T_z and T_θ indicate extension. Since the electrical excitation is axisymmetric all quantities are independent of angular position.

The two-dimensional electroelastic equations of state can be expressed as

$$S_z(z,\omega) = s_{12}^E T_\theta(z,\omega) + s_{11}^E T_z(z,\omega) + d_{31} E_r(\omega), \tag{2a}$$

$$S_\theta(z,\omega) = s_{11}^E T_\theta(z,\omega) + s_{12}^E T_z(z,\omega) + d_{31} E_r(\omega), \tag{2b}$$

and

$$D_r(z,\omega) = d_{31}[T_\theta(z,\omega) + T_z(z,\omega)] + \epsilon_{33}^T E_r(\omega), \tag{2c}$$

where S_z and S_θ are the strains in the axial and tangential directions, respectively, and D_r is the charge density. The electric field E_r is independent of z since the outer and inner surfaces of the cylinder are electroded. The normal stress in the radial direction, T_r, has been assumed to be zero since the cylinder is thin and there are no external loads on its curved surfaces. The electroelastic coefficients s_{11}^E, s_{12}^E, d_{31}, and ϵ_{33}^T have their usual definitions.

The strain-displacement relationships for cylindrical membranes can be expressed as

$$S_z(z,\omega) = \partial U(z,\omega)/\partial z \tag{3a}$$

and

$$S_\theta(z,\omega) = W(z,\omega)/a. \tag{3b}$$

It can be easily seen by first rearranging Eqs. (2a) and (2b) and using Eq. (3) that

$$T_z(z,\omega) = [Y/(1-\sigma^2)][\partial U(z,\omega)/\partial z + \sigma W(z,\omega)/a$$
$$- d_{31}(1+\sigma)E_r(\omega)] \tag{4a}$$

and

$$T_\theta(z,\omega) = [Y/(1-\sigma^2)][\sigma\partial U(z,\omega)/\partial z + W(z,\omega)/a$$
$$- d_{31}(1+\sigma)E_r(\omega)], \tag{4b}$$

where

$$Y = 1/s_{11}^E \tag{4c}$$

and

$$\sigma = -s_{12}^E/s_{11}^E. \tag{4d}$$

The equations of motion can then be obtained by eliminating T_z and T_θ from Eqs. (1) and (4) and expressed in the form

$$[\Gamma]\begin{Bmatrix} U(z,\omega) \\ W(z,\omega) \end{Bmatrix} - \omega^2\rho h\begin{Bmatrix} U(z,\omega) \\ W(z,\omega) \end{Bmatrix} = \begin{Bmatrix} 0 \\ CV(\omega) \end{Bmatrix}; \tag{5a}$$

where

$$[\Gamma] = \frac{Yh}{a^2(1-\sigma^2)}\begin{bmatrix} -a^2\partial^2/\partial z^2 & -\sigma a\partial/\partial z \\ \sigma a\partial/\partial z & 1 \end{bmatrix}, \tag{5b}$$

$$C = d_{31}Y/[a(1-\sigma)], \tag{5c}$$

and

$$V(\omega) = E_r(\omega)h \tag{5d}$$

is the applied voltage.

The boundary conditions can be expressed in the form

$$[B]\begin{Bmatrix} U(z,\omega) \\ W(z,\omega) \\ V(\omega) \end{Bmatrix} = 0, \tag{6}$$

where $[B]$ is a 1×3 matrix whose elements are either constants or differential operators.

It is assumed that $V(\omega)$ is equal to a constant V_0 in the rest of this article.

B. Solution: Displacement and input electrical admittance

The solution to the inhomogeneous equations of motion (5) which satisfies the boundary conditions in Eq. (6) can be expressed in the form

$$\begin{Bmatrix} U(z,\omega) \\ W(z,\omega) \end{Bmatrix} = \begin{Bmatrix} U_{dc}(z) \\ W_{dc}(z) \end{Bmatrix} + \begin{Bmatrix} U_{sc}(z,\omega) \\ W_{sc}(z,\omega) \end{Bmatrix}. \tag{7}$$

The first term on the right-hand side of Eq. (7) is the dc solution, i.e., it satisfies the static equations

$$[\Gamma]\begin{Bmatrix} U_{dc}(z) \\ W_{dc}(z) \end{Bmatrix} = \begin{Bmatrix} 0 \\ CV_0 \end{Bmatrix} \tag{8}$$

obtained by setting $\omega=0$ in Eqs. (5), and also satisfies two boundary conditions[5] which can be expressed in the form

$$[B]\begin{Bmatrix} U_{dc}(z) \\ W_{dc}(z) \\ V_0 \end{Bmatrix} = 0. \tag{9}$$

It can be easily seen that the solution to Eq. (8) can be expressed as

$$\begin{Bmatrix} U_{dc}(z) \\ W_{dc}(z) \end{Bmatrix} = V_0 \begin{Bmatrix} U_0 + U_1 z \\ W_0 \end{Bmatrix}, \tag{10a}$$

where

$$W_0 = d_{31} a (1+\sigma)/h - \sigma a U_1 \tag{10b}$$

and the constants U_0 and U_1 are to be determined by using the boundary conditions in Eq. (9).

The second term in the right-hand side of Eq. (7) can be expressed as

$$\begin{Bmatrix} U_{sc}(z,\omega) \\ W_{sc}(z,\omega) \end{Bmatrix} = V_0 \sum_{m=M}^{M} \sum_{i=1}^{I_m} \mathbb{V}_{mi}(\omega)\{\Phi_{mi}(z)\}, \tag{11a}$$

where

$$\{\Phi_{mi}(z)\} = \begin{Bmatrix} \psi_{mi}(z) \\ \phi_{mi}(z) \end{Bmatrix} \tag{11b}$$

are the eigenfunctions of the short-circuited cylinder, $M=0$ or 1, $M=\infty$, and $I_m=1$ or 2. Each eigenfunction satisfies the homogeneous equations for the short-circuited cylinder,

$$[\Gamma]\begin{Bmatrix} U_{sc}(z,\omega) \\ W_{sc}(z,\omega) \end{Bmatrix} - \omega^2 \rho h \begin{Bmatrix} U_{sc}(z,\omega) \\ W_{sc}(z,\omega) \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}, \tag{12}$$

at an angular resonance frequency ω_{mi}, and the boundary conditions

$$[B]\begin{Bmatrix} U_{sc}(z,\omega) \\ W_{sc}(z,\omega) \\ 0 \end{Bmatrix} = 0 \tag{13}$$

at all frequencies.

It is noted that the eigenfunctions form a complete set[7] of functions. However, in most cases, they are complete only in the normwise sense and not in the pointwise sense.[8] Sets of functions which are complete in the ordinary Fourier or normwise sense and are not all zero at the same point are said to be pointwise complete.

The displacement coefficients per unit voltage, $\mathbb{V}_{mi}(\omega)$, can be determined by their orthogonal property, as shown later.

The dc solution is included in the right-hand side of Eq. (7) for two reasons. First, in general, the boundary conditions have the form of Eq. (6) and depend on V_0. However, the eigenfunctions satisfy only the boundary conditions of the short-circuited cylinder, i.e., the boundary conditions when $V_0=0$. Therefore, an additional term is required to ensure that the proper boundary conditions are satisfied. It can be easily verified by substituting Eq. (7) into Eq. (6) and using Eqs. (9) and (13) that the boundary conditions of the electrically driven piezoelectric cylinder are satisfied. The second reason is related to the first. It can be seen that the additional term need not be the dc solution. In fact, it can be any term which satisfies the boundary conditions in Eq. (9) and need not satisfy the static equations (8). However, if an arbitrary term is used to satisfy the boundary conditions, it will not have any physical significance. Further, the solution converges to an incorrect displacement distribution in some

cases if an arbitrary term is used. If the dc solution is used, it not only has physical significance but rapid convergence to the correct displacement distribution can be expected, especially at low frequencies.

Next, the displacement coefficients per unit voltage, $\mathbb{V}_{mi}(\omega)$, can be determined by using the equations of motion of the piezoelectric cylinder and the orthogonal property of the eigenfunctions. By definition

$$2\pi a \int_0^L \{\Phi_{nk}(z)\}^T \{\Phi_{mi}(z)\}dz = N_{mi}\delta_{mn}\delta_{ik}, \tag{14a}$$

where N_{mi} are constants and δ_{mn} is the Kronecker delta. Further,

$$[\Gamma]\{\Phi_{mi}(z)\} = \omega_{mi}^2 \rho h \{\Phi_{mi}(z)\}, \tag{14b}$$

where ω_{mi} is the angular resonance frequency at which $\{\Phi_{mi}(z)\}$ is a solution to the homogeneous equations of motion (12). On premultiplying Eq. (14b) by $\{\Phi_{nk}(z)\}^T$, integrating over the mean surface of the shell, and using the orthogonal property in Eq. (14a), it can be seen that

$$2\pi a \int_0^L \{\Phi_{nk}(z)\}^T [\Gamma]\{\Phi_{mi}(z)\}dz = \omega_{mi}^2 \rho h N_{mi}\delta_{mn}\delta_{ik}. \tag{14c}$$

It can now be seen by substituting the form of the solution in Eq. (7) into the inhomogeneous governing Eqs. (5), premultiplying them by $\{\Phi_{nk}(z)\}^T$ after using Eq. (8), integrating over the mean surface of the shell, and using Eqs. (14a)–(14c) that

$$V_0 \mathbb{V}_{mi}(\omega) = \frac{F_{mi}(\omega)}{K_{mi} - M_{mi}\omega^2}, \tag{15a}$$

where

$$F_{mi}(\omega) = \omega^2 2\pi a\rho h \int_0^L \{\Phi_{mi}(z)\}^T \begin{Bmatrix} U_{dc}(z) \\ W_{dc}(z) \end{Bmatrix} dz, \tag{15b}$$

$$K_{mi} = \omega_{mi}^2 \rho h N_{mi}, \tag{15c}$$

and

$$M_{mi} = \rho h N_{mi}. \tag{15d}$$

It is noted that $F_{mi}(\omega)$, $K_{mi}(\omega)$, and $M_{mi}(\omega)$ are the modal force, stiffness, and mass, respectively. It is also noted that F_{mi} is zero at zero frequency, as indeed should be the case.

The input electrical admittance can now be obtained. After substituting Eqs. (4) into the equation of state (2c) and integrating over the mean surface, the total charge on the mean surface, $Q(\omega)$, can be expressed as

$$Q(\omega)=2\pi a\int_0^L D_r(z,\omega)dz$$

$$=2\pi a\left[E_r\epsilon_{33}^T L[1-2k_{31}^2/(1-\sigma)]\right.$$

$$+\frac{d_{31}}{s_{11}^E(1-\sigma)}\left(U(L,\omega)-U(0,\omega)\right.$$

$$\left.\left.+\int_0^L\frac{W(z,\omega)}{a}dz\right)\right], \tag{16a}$$

where

$$k_{31}^2=d_{31}^2/(s_{11}^E\epsilon_{33}^T). \tag{16b}$$

Finally, after using Eqs. (14) and (9), the input electrical admittance, $Y(\omega)$, can be expressed as

$$Y(\omega)=j\omega Q(\omega)/V_0 \tag{17a}$$

$$=j\omega C_0\left\{1-k_{31}^2+(d_{31}h/[\epsilon_{33}^T s_{11}^E(1-\sigma)L])\right.$$

$$\times\left[U_1 L(1-\sigma)+\sum_{m=1}^M\sum_{i=1}^2\mathbb{V}_{mi}(\omega)\left(\psi_{mi}(L)-\psi_{mi}(0)\right.\right.$$

$$\left.\left.\left.+\frac{1}{a}\int_0^L\phi_{mi}(z)dz\right)\right]\right\}, \tag{17b}$$

where

$$C_0=2\pi aL\epsilon_{33}^T/h, \tag{17c}$$

and $M=\infty$.

II. SPECIAL CASES

Four special cases are presented to illustrate the approach. They are a cylinder with zero axial stress at both ends, a cylinder with one end axially fixed and the other with zero axial stress, a cylinder radially fixed at both ends, and finally, a cylinder axially fixed at both ends.

A. Free–free cylindrical membrane

The first special case consider a cylindrical membrane with zero axial stress at the boundaries. The boundary conditions of the electrically driven cylinder are

$$-a^2\frac{\partial}{\partial z}U(z,\omega)-\sigma aW(z,\omega)+d_{31}a^2(1+\sigma)V_0/h=0$$

$$\text{at } z=0,L. \tag{18a}$$

Under static conditions, Eqs. (18a) reduce to

$$-a^2\frac{\partial}{\partial z}U_{dc}(z)-\sigma aW_{dc}(z)+d_{31}a^2(1+\sigma)V_0/h=0$$

$$\text{at } z=0,L. \tag{18b}$$

The static displacement can be easily obtained by using Eqs. (10a), (10b), and (18b). It can be seen that

$$U_1=d_{31}/h \tag{19a}$$

and

$$W_0=d_{31}a/h. \tag{19b}$$

Since the boundary conditions permit rigid translation of the shell, U_0 cannot be determined by using the boundary conditions. However, since the excitation is symmetric with respect to the center of the cylinder, rigid translation is not possible and the axial displacement must be antisymmetric with respect to the center. It therefore follows that the dc solution can be expressed as

$$\begin{Bmatrix}U_{dc}(z)\\W_{dc}(z)\end{Bmatrix}=V(\omega)\begin{Bmatrix}d_{31}(z-L/2)/h\\d_{31}a/h\end{Bmatrix}. \tag{20}$$

The short-circuit boundary conditions are

$$-a^2\frac{\partial}{\partial z}U_{sc}(z,\omega)-\sigma aW_{sc}(z,\omega)=0 \quad \text{at } z=0,L. \tag{21}$$

It can be seen from Eqs. (12) and (21) that the eigenfunctions of the short-circuited cylinder can be expressed in the form

$$\{\Phi_{mi}(z)\}=\begin{Bmatrix}C_{mi}\cos(k_m z)\\\sin(k_m z)\end{Bmatrix}, \tag{22}$$

where $m=1,2,3,...,$ and $i=1,2$. The boundary conditions in Eq. (21) are satisfied when $k_m=m\pi/L$. The eigenfunctions are not pointwise complete since $\sin(k_m z)$ is zero at $z=0$ and L for all m. Since the eigenfunctions must satisfy the homogenous equations of motion (12) also, it follows that

$$\begin{bmatrix}k_m^2 a^2 & -\sigma k_m a\\-\sigma k_m a & 1\end{bmatrix}\begin{Bmatrix}C_{mi}\\1\end{Bmatrix}-\Omega^2\begin{Bmatrix}C_{mi}\\1\end{Bmatrix}=\begin{Bmatrix}0\\0\end{Bmatrix}, \tag{23a}$$

where

$$\Omega^2=\omega^2 a^2\rho(1-\sigma^2)/Y. \tag{23b}$$

A solution to Eq. (23a) exists only when

$$\begin{vmatrix}k_m^2 a^2-\Omega^2 & -\sigma k_m a\\-\sigma k_m a & 1-\Omega^2\end{vmatrix}=0. \tag{24}$$

After solving the above quadratic equation to determine Ω_{mi}, it can be seen by using Eq. (23a) that

$$C_{mi}=\sigma k_m a/(k_m^2 a^2-\Omega_{mi}^2)=(1-\Omega_{mi}^2)/\sigma k_m a. \tag{25}$$

Now, after using Eqs. (14a) and (22) to obtain

$$N_{mi}=\pi aL(1+C_{mi}^2), \tag{26}$$

it can be shown by using Eqs. (15), (20), (24), and (26) that

$$\mathbb{V}_{mi}(\omega)=\omega^2 4d_{31}\frac{(k_m a-C_{mi})}{k_m Lk_m h(1+C_{mi}^2)(\omega_{mi}^2-\omega^2)}$$

$$m=1,3,5,...$$

$$=0 \quad m=2,4,6,..., \tag{27a}$$

where

$$\omega_{mi}^2=\Omega_{mi}^2 Y/[a^2\rho(1-\sigma^2)]. \tag{27b}$$

It then follows from Eqs. (15), (17), (20), and (22) that the input electrical admittance can be expressed as

$$Y(\omega) = j\omega C_0 \left\{ 1 + (2d_{31}h/[\epsilon_{33}^T s_{11}^E(1-\sigma)L]) \right.$$
$$\left. \times \left(\sum_{m=1}^{M} \sum_{i=1}^{2} \mathbb{V}_{mi}(\omega) \left[\frac{1}{k_m a} - C_{mi} \right] \right) \right\}. \quad (28)$$

B. Fixed-free cylindrical membrane

As the second special case consider a cylindrical membrane axially fixed at $z=0$ and with no axial stress at $z=L$. The boundary conditions can be expressed as

$$U(z,\omega) = 0, \quad z = 0 \quad (29a)$$

and

$$-a^2 \frac{\partial}{\partial z} U(z,\omega) - \sigma a W(z,\omega) + d_{31}a^2(1+\sigma)V_0/h = 0,$$
$$z = L. \quad (29b)$$

In contrast to the first special case, the boundary conditions are not the same at the two ends. Therefore, the axial and radial displacements will not be antisymmetric and symmetric, respectively, with respect to the center of the cylinder. However, Eqs. (29a) and (29b) are satisfied at the center and ends, respectively, of the cylinder considered in the first special case. Therefore, some of the following equations can be inferred from the analysis of the first special case.

The static boundary conditions can be obtained by replacing $U(z,\omega)$ and $W(z,\omega)$ in Eqs. (29) by $U_{dc}(z)$ and $W_{dc}(z)$, respectively. Then, it can be shown by using Eqs. (10) that

$$\begin{Bmatrix} U_{dc}(z) \\ W_{dc}(z) \end{Bmatrix} = V_0 \begin{Bmatrix} d_{31}z/h \\ d_{31}a/h \end{Bmatrix}. \quad (30)$$

The short-circuit boundary conditions can be expressed as

$$U_{sc}(z,\omega) = 0, \quad z = 0 \quad (31a)$$

and

$$-a^2 \frac{\partial}{\partial z} U_{sc}(z,\omega) - \sigma a W_{sc}(z,\omega) = 0, \quad z = L. \quad (31b)$$

Equations (10) and (31) can then be used to show that the eigenfunctions for this special case are

$$\{\Phi_{mi}(z)\} = \begin{Bmatrix} C_{mi} \sin(k_m z) \\ \cos(k_m z) \end{Bmatrix}, \quad (32a)$$

where $m = 1,2,3,...$, $i = 1,2$, $k_m = (2m-1)\pi/(2L)$,

$$C_{mi} = -\sigma k_m a/(k_m^2 a^2 - \Omega_{mi}^2) = -(1 - \Omega_{mi}^2)/\sigma k_m a,$$
$$(32b)$$

and Ω_{mi}^2 are the solutions to

$$\begin{vmatrix} k_m^2 a^2 - \Omega^2 & \sigma k_m a \\ \sigma k_m a & 1 - \Omega^2 \end{vmatrix} = 0. \quad (32c)$$

The displacement coefficients per unit voltage can now be obtained by using Eqs. (15), (30), and (32). It can be shown that

$$\mathbb{V}_{mi}(\omega) = \omega^2 2(-1)^{m+1} d_{31}$$
$$\times \frac{(k_m a + C_{mi})}{k_m L k_m h(1 + C_{mi}^2)(\omega_{mi}^2 - \omega^2)}. \quad (33)$$

Finally, after using Eqs. (17), (30), and (32), the input electrical admittance can be expressed as

$$Y(\omega) = j\omega C_0 \left[1 + (d_{31}h/[\epsilon_{33}^T s_{11}^E(1-\sigma)L]) \left(\sum_{m=1}^{M} (-1)^{m+1} \right. \right.$$
$$\left. \left. \times \sum_{i=1}^{2} \mathbb{V}_{mi}(\omega) \left[C_{mi} + \frac{1}{k_m a} \right] \right) \right]. \quad (34)$$

C. Radially fixed–fixed cylindrical membrane

As the third special case, consider a cylinder with radially fixed boundaries, i.e.,

$$W(z) = 0, \quad z = 0, L. \quad (35)$$

After using the static boundary conditions and Eq. (10b), it can be shown that

$$W_0 = 0 \quad (36a)$$

and

$$U_1 = d_{31}(1+\sigma)/(\sigma h). \quad (36b)$$

U_0 cannot be determined directly from the boundary conditions because rigid translation is possible. However, since the excitation is purely radial, the axial displacement must be symmetric about the center of the cylinder. It therefore follows that

$$\begin{Bmatrix} U_{dc}(z) \\ W_{dc}(z) \end{Bmatrix} = V_0(\omega) \begin{Bmatrix} d_{31}(1+\sigma)(z-L/2)/(\sigma h) \\ 0 \end{Bmatrix}. \quad (37)$$

It can be easily shown that the eigenfunctions of the short-circuited cylinder are given by Eq. (22) where $m = 1,2,3,...$, $i = 1,2$, $k_m = m\pi/L$, and that the values of C_{mi} are given by Eq. (25). The angular resonance frequencies can be obtained from Eq. (24). Even though the eigenfunctions and the resonance frequencies for this special case are the same as those for the first special case, the response is not the same.

It can be shown after using Eqs. (15) and (22) that

$$\mathbb{V}_{mi}(\omega) = -\omega^2 4 d_{31}(1+\sigma)/\sigma$$
$$\times \frac{C_{mi}}{k_m L k_m h(1 + C_{mi}^2)(\omega_{mi}^2 - \omega^2)} \quad m = 1,3,5,...$$
$$= 0 \quad m = 2,4,6,... . \quad (38)$$

Finally, it can be shown by using Eqs. (17) and (37) that the input electrical admittance can be expressed as

$$Y(\omega) = j\omega C_0 \left\{ 1 + \frac{k_{31}^2}{\sigma(1-\sigma)} + (2d_{31}h/[\epsilon_{33}^T s_{11}^E(1 \right.$$
$$\left. -\sigma)L]) \left(\sum_{m=1}^{M} \sum_{i=1}^{2} \mathbb{V}_{mi}(\omega) \left[\frac{1}{k_m a} - C_{mi} \right] \right) \right\}. \quad (39)$$
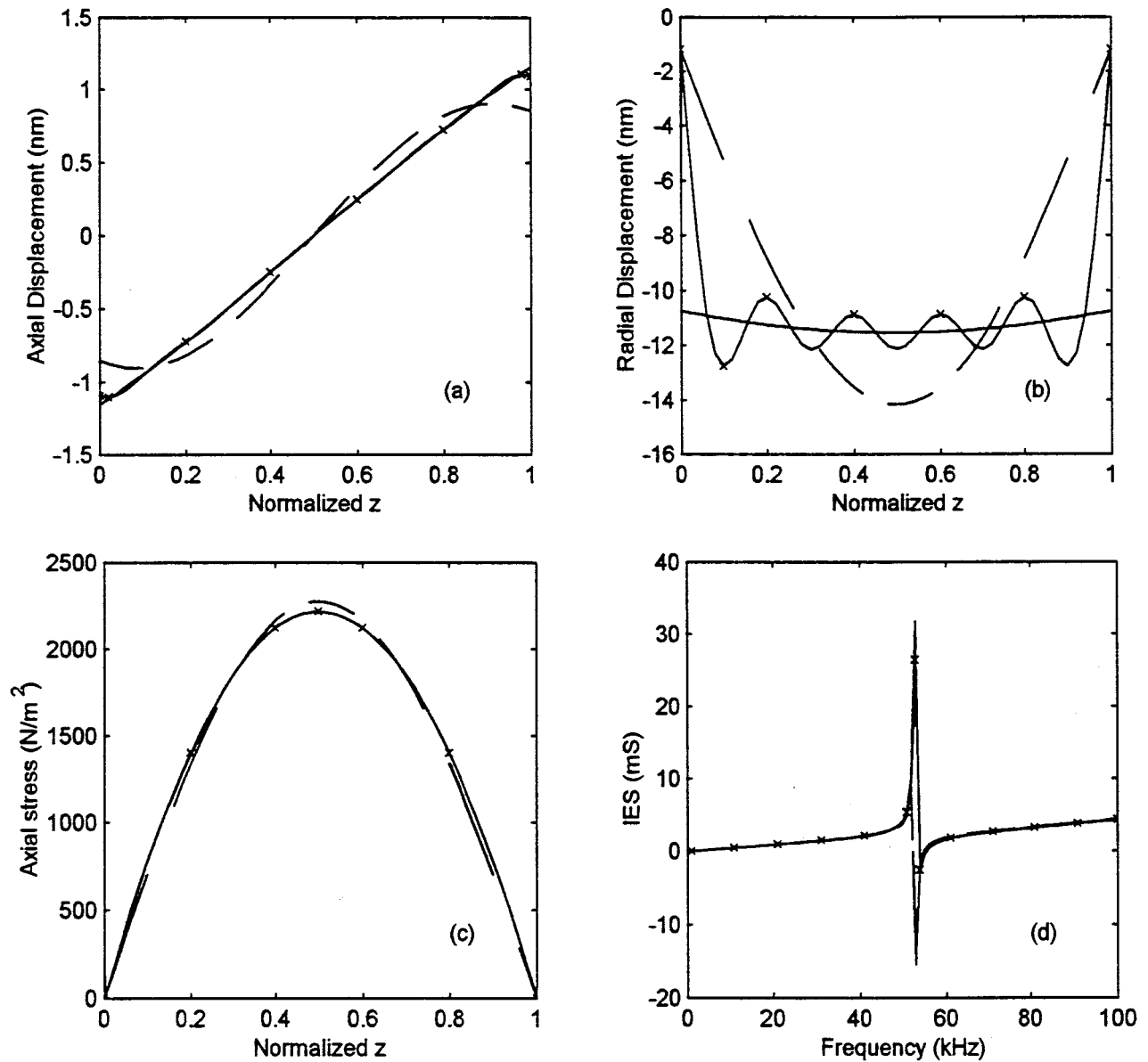
FIG. 2. The (a) axial displacement, (b) radial displacement, and (c) axial stress when 1 V is applied at 50 kHz, and (d) the input electrical susceptance of a cylinder with zero axial stress at both ends. —— closed form; ---- eigenfunction approach, $M=1$; —✕— eigenfunction approach with five nonzero m terms, i.e., $M=9$.

It is shown in the Appendix that an incorrect solution is obtained if the dc solution is neglected and it is assumed that the displacement can expressed as a weighted sum of eigenfunctions only.

D. Axially fixed–fixed cylindrical membrane

As the fourth special case consider a cylindrical shell with axially fixed boundaries, i.e.,

$$U(z,\omega)=0, \quad z=0,L. \tag{40}$$

For the static case, it can be easily seen by using the general form of the solution in Eqs. (10) and the boundary conditions

$$U_{\mathrm{dc}}(z)=0, \quad z=0,L \tag{41}$$

that

$$\begin{Bmatrix} U_{\mathrm{dc}}(z) \\ W_{\mathrm{dc}}(z) \end{Bmatrix} = V_0 \begin{Bmatrix} 0 \\ d_{31}a(1+\sigma)/h \end{Bmatrix}. \tag{42}$$

All except one of the eigenfunctions of the short-circuited cylinder have the form shown in Eq. (32a), where $m=1,2,3,...,$ $i=1,2,$ $k_m=m\pi/(2L)$, and C_{mi} and Ω_{mi} can be obtained from Eqs. (32b) and (32c), respectively.

It can be easily seen that

$$\{\Phi_{mi}(z)\}=\begin{Bmatrix} 0 \\ 1 \end{Bmatrix}, \tag{43}$$

where $m=0$ and $i=1$ is also an eigenfunction which satisfies the homogeneous governing equation at $\Omega=1$, i.e., $\Omega_{01}=1$.

It is noted that ϕ_{mi} $m=1,2,3,...$ and the electrical excitation are antisymmetric and symmetric, respectively, with respect to the center of the cylinder. It therefore follows from
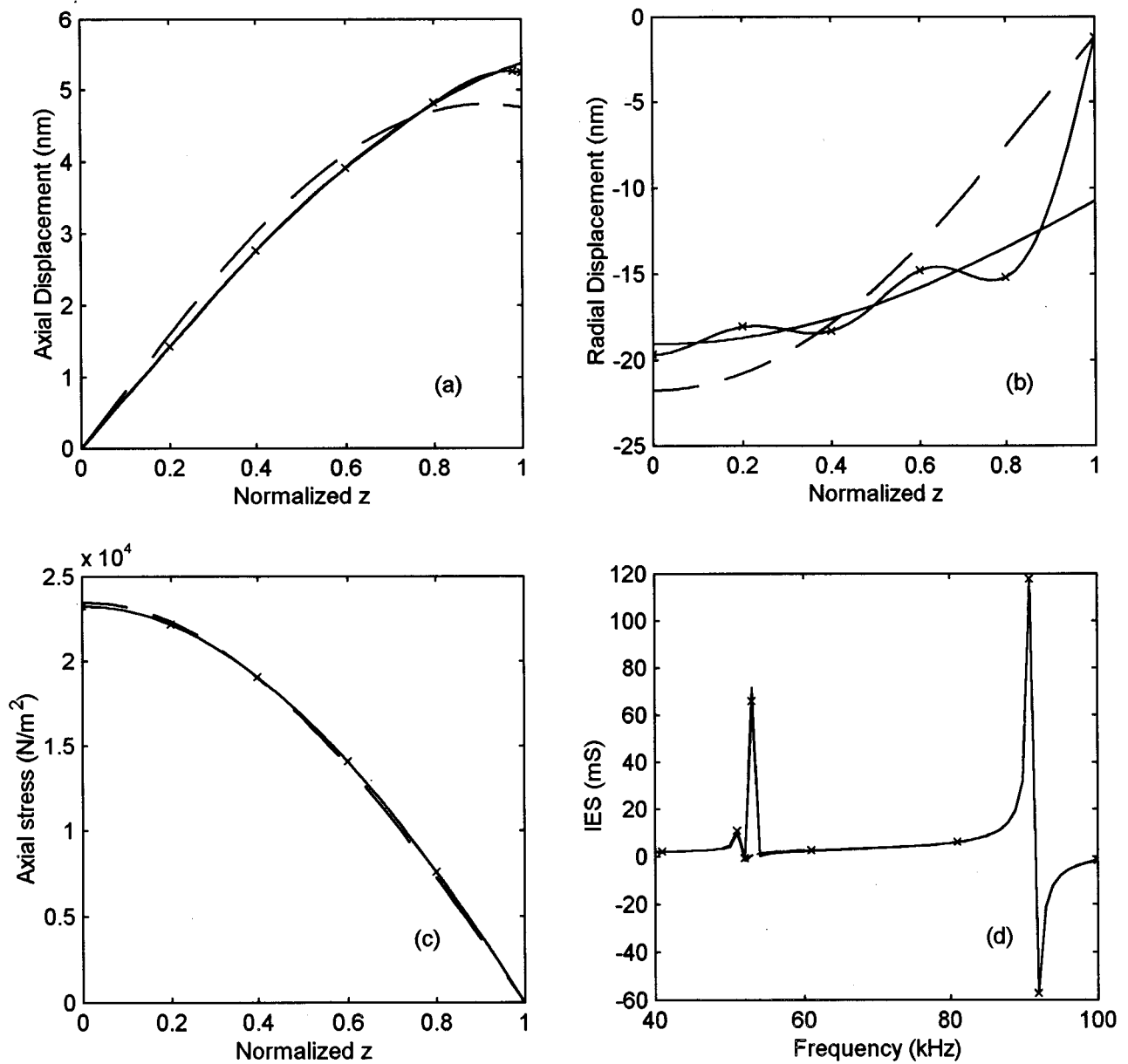
FIG. 3. The (a) axial displacement, (b) radial displacement, and (c) axial stress when 1 V is applied at 50 kHz, and (d) the input electrical susceptance of a cylinder axially fixed at the left end and with zero axial stress at the right end. —— closed form; ---- eigenfunction approach, $M=1$; —×— eigenfunction approach with five nonzero m terms, i.e., $M=5$.

Eqs. (15) and (42) that $\mathbb{V}_{mi}$ are zero for these values of m. However, when $m=0$, it can be shown by using Eqs. (15) and (43) and $N_{01}=2\pi aL$ that

$$\mathbb{V}_{01}(\omega)=\omega^2 d_{31}a(1+\sigma)/h\,\frac{1}{(\omega_{01}^2-\omega^2)}.\qquad(44)$$

It therefore follows from Eqs. (7), (11a), (42), and (44) that

$$\begin{Bmatrix} U(z,\omega) \\ W(z,\omega) \end{Bmatrix}=V_0 d_{31}a(1+\sigma)/h\begin{Bmatrix} 0 \\ 1/(1-\Omega^2) \end{Bmatrix}.\qquad(45)$$

Finally, it can be seen from Eqs. (16), (17a), and (45) that

$$Y(\omega)=j\omega C_0\left\{1-\frac{k_{31}^2}{1-\sigma}\left(2-\frac{1+\sigma}{1-\Omega^2}\right)\right\}.\qquad(46)$$

The expressions for the displacement distribution and input electrical admittance in Eqs. (45) and (46), respectively, are in exact agreement with the results obtained by using Ref. 5.

III. NUMERICAL RESULTS AND DISCUSSION

The input electrical admittances of cylinders with various boundary conditions computed using the eigenfunction approach are compared with those computed using closed form expressions. The following parameters are used: length$=10$ mm, $a=10$ mm, $h=1$ mm, $\epsilon_{33}^T/\epsilon_0=1300$, $s_{11}^E=12\times10^{-12}$ m^2/N, $s_{12}^E=-4\times10^{-12}$ m^2/N, $d_{31}=-120\times10^{-12}$ C/N, and $\rho=7500$ kg/m^3. Since all the electroelastic coefficients used are real, the displacements go to infinity
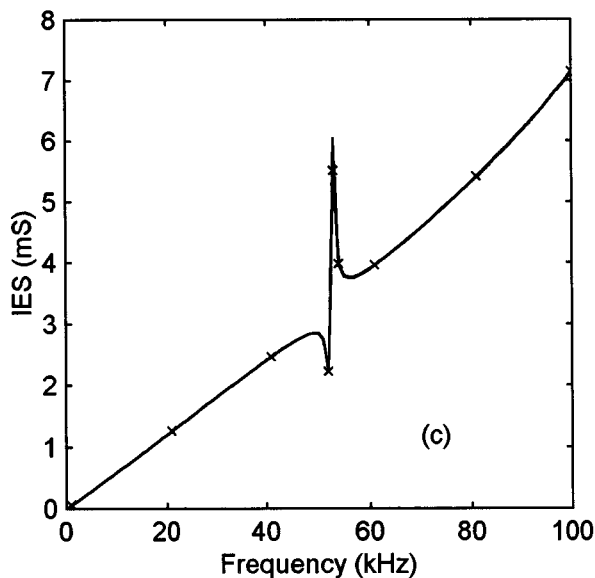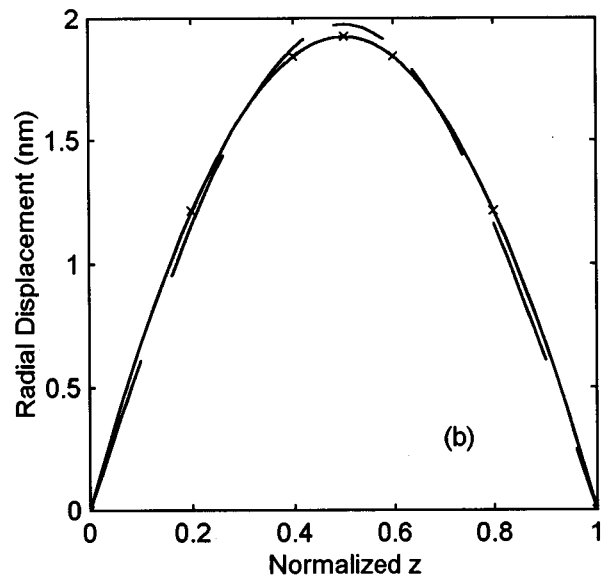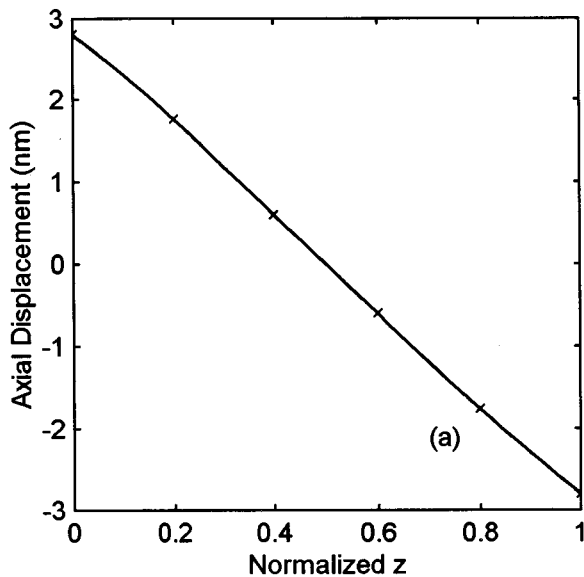
FIG. 4. The (a) axial displacement and (b) radial displacement when 1 V is applied at 50 kHz and (c) the input electrical susceptance of a cylinder radially fixed at both ends. —— closed form; ---- eigenfunction approach, $M=1$; —×— eigenfunction approach with five nonzero m terms, i.e., $M = 9$.

at resonance and the input electrical admittance is purely imaginary. Only the imaginary part, the input electrical susceptance (IES), is shown in the figures.

The closed form solution obtained by using Ref. 5 is shown by a solid line in each of the figures. Results obtained by using the eigenfunction approach and $M = 1$ in Eqs. (11a) and (17b) are shown by a dashed line. A solid-× line is used to show results obtained using five nonzero terms in the summation over m, i.e., $M = 9$ for the first and third special cases and $M = 5$ for the second special case. It is not possible to distinguish between these lines in some of the figures. All the displacement and stress distributions are shown when 1 V is applied at a frequency of 50 kHz, which is a little less than the lowest resonance frequency in each special case. All displacements are shown in nanometers.

When the membrane approximation is used, all the resonances in the lower branch, which are infinite in number, occur below the hoop mode resonance. However, all of them cannot be seen in the input electrical susceptances shown in the figures since a resolution of only 1 kHz has been used.

The hoop mode resonance frequency, $(Y/\rho)^{1/2}/(2\pi a)$, is equal to 53.05 kHz for the cylinder used here.

Numerical results for the first special case are shown in Fig. 2(a)–(d). The axial and radial displacements and the axial stress are shown in Fig. 2(a), (b), and (c), respectively. The lowest resonance frequency, $f_{11} = \omega_{11}/(2\pi)$, in this case is equal to 52.8 kHz.

It can be seen from Fig. 2(a) that the agreement between the closed form and the eigenfunction solutions is quite good even when $M = 1$ is used. When $M = 9$ is used, the closed form and eigenfunction solutions are almost identical except near the ends of the cylinder. It is noted that the slope of the axial displacement at the ends of the cylinder does not depend on M. This can also be seen by using Eq. (22).

It can be seen from Fig. 2(a) and (b) that the difference between the radial displacements obtained by using the closed form and eigenfunction approaches is more than the difference between the axial displacements. It can also be seen from Fig. 2(b) that the displacements at the ends of the cylinder obtained by using the present approach are indepen-
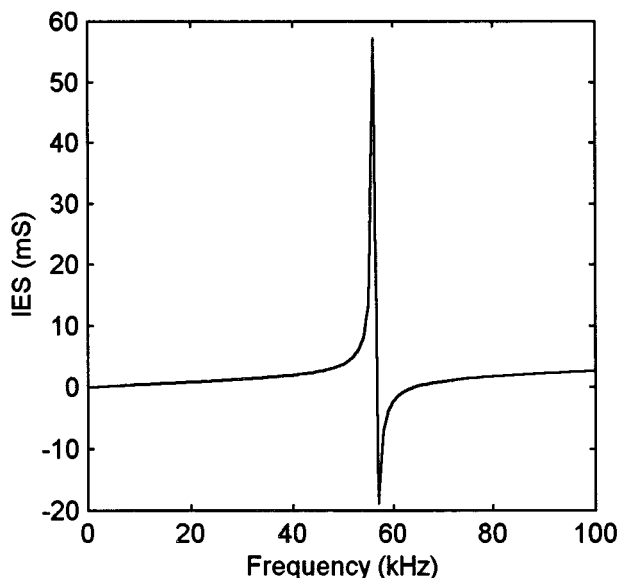
FIG. 5. The input electrical susceptance of a cylinder axially fixed at both ends.

dent of M and equal to the static displacements. This occurs because all the radial components of the eigenfunctions for this case are zero at the ends of the cylinder, i.e., they form only a normwise and not a pointwise complete set of functions. When $M=9$, the eigenfunction solution oscillates about the closed form solution. The magnitude of the oscillations near the center of the cylinder decreases when a larger value of M is used. However, the values at the ends remain the same. The correct values at the ends can be approximately determined by using a large value of M and extrapolation.

The axial stress is shown in Fig. 2(c). At the ends of the cylinder the stresses obtained using the closed form and eigenfunction approaches are equal. The slopes of the stresses are also equal. Thus the boundary conditions are exactly satisfied even though the eigenfunctions are complete only in the normwise sense.

The input electrical susceptances are shown in Fig. 2(d). The agreement between the closed form and eigenfunction approaches is very good, except near the hoop mode resonance frequency, even when $M=1$. The agreement near this frequency is much better when $M=9$. This occurs because none of the terms in the eigenfunction approach resonate at this frequency. However, the sum of the terms goes to infinity. Theoretical and experimental values of IES for this special case have been compared in Ref. 9.

Numerical results for the second special case, a cylinder axially fixed at $z=0$ and axially free at $z=L$, are shown in Fig. 3(a)–(d). The lowest resonance frequency, f_{11}, in this case is 51.3 kHz. It can be seen from Fig. 3(a) and (c) that the boundary conditions are exactly satisfied. It is noted that the maximum axial stress in this case is about 10 times the maximum stress in the previous case. This is primarily because the drive frequency is closer to a resonance frequency. The resonance at about 90 kHz seen in Fig. 3(d) is f_{12}.

Results for the third special case are shown in Fig. 4(a)–(c). The cylinder has zero radial displacement at both ends.

The boundary conditions are again exactly satisfied. In this case extremely good agreement is obtained even when $M=1$. The lowest resonance frequency is 52.8 kHz.

The input electrical susceptance of a cylinder with zero axial displacement at both ends is shown in Fig. 5. Only one curve is shown since the expression for the admittance obtained in this article is the same as the one obtained earlier in Ref. 5. It is noted that there is only one resonance frequency in this case.

IV. CONCLUSIONS

An eigenfunction approach to determine the input electrical admittance and the displacement distribution of a radially polarized plezoelectric cylindrical shells of finite length has been presented. In this approach the equations of motion and the boundary conditions obtained using the membrane approximation are exactly satisfied. The approach is based on the use of the eigenfunctions of the short-circuited cylinder and the solution to the static problem. It was shown that the series expressions for the displacement and stress distributions and the input electrical admittance converge very rapidly to the closed form solutions. It was also shown, through an example, that the series may converge to a wrong displacement distribution if the static solution is not used. McGee et al.[10] have shown that the resonance frequencies of plates obtained by using a Ritz method are incorrect in some cases if the static solution is not used in conjunction with a complete set of functions. The approach can be used to solve other problems for which closed form solutions are not known.

ACKNOWLEDGMENT

APPENDIX: A CASE OF CONVERGENCE TO A WRONG SOLUTION

Consider, again, the radially fixed–fixed cylinder. The boundary conditions are given by Eq. (35). The static or dc solution is not required to satisfy the boundary conditions. It is therefore assumed that the displacement distribution can be expressed as

$$\begin{Bmatrix} U(z,\omega) \\ W(z,\omega) \end{Bmatrix} = V_0 \sum_{m=1,3,5}^{\infty} \sum_{i=1}^{2} \mathbb{V}_{mi}(\omega)\{\Phi_{mi}(z)\}, \tag{A1}$$

where the eigenfunctions $\{\Phi_{mi}(z)\}$ satisfy the homogeneous form of Eq. (5a) and the boundary conditions in Eq. (35), and have the form shown in Eq. (22). The values of k_m, Ω_{mi}, and C_{mi} are the same as those in Sec. II C. The eigenfunctions are complete only in the normwise sense and are not complete in the pointwise sense since $\sin(k_m z)$ is zero at $z=0$ and L for all m.

It is clear that the values of U and W obtained by using Eq. (A1) are incorrect because the radial displacement distribution is discontinuous at the ends of the cylinder. The radial displacement is discontinuous for the axially free–free and axially fixed-free cases also as seen from Figs. 2(b) and 3(b),

respectively. However, in those cases, the radial displacement is not specified at the boundaries and therefore the radial displacement is correct, except at the ends. The correct radial displacement at the ends can be determined by extrapolation. In this case, the radial displacement must be zero at the boundaries and the boundary conditions are satisfied only because of the discontinuity. Extrapolation will lead to $W(z,\omega) \neq 0$ at the ends and the boundary conditions will not be satisfied. The difficulty arises because the eigenfunctions do not form a pointwise complete set of functions.

The radial displacement obtained by assuming that displacement distribution has the form in Eq. (7) instead of the form in Eq. (A1) does not have any discontinuities at the ends as seen in Fig. 4(b). Further, it satisfies the boundary conditions. Therefore, the displacement obtained by using Eq. (7) is the correct solution.

It is noted that it can be determined whether or not the solution obtained by using the eigenfunction approach is correct by examining it. It is not necessary to compare the eigenfunction solution with the solution obtained by some other approach to determine if it is correct. Further, it has been shown that the correct solution can be obtained for all the cases considered here.

It is also noted that both the eigenfunction solutions obtained using Eq. (A1) and Eq. (7), respectively, satisfy the equations of motion (5) in a least-squares sense. U and W, computed using Eq. (A1), satisfy the axial equation of motion at $0 \leq z \leq L$ and the radial equation of motion at $0 < z < L$. However, U and W, computed using Eq. (7), satisfy the axial and radial equations of motion at $0 < z < L$ and at $0 \leq z \leq L$, respectively.

When $\omega \ll \omega_{mi}$, it can be seen, through numerical computation, that the series in Eq. (A1) converges to

$$U(z,\omega) = A(z - L/2) \quad 0 \leq z \leq L \tag{A2}$$

and

$$W(z,\omega) = B \quad 0 < z < L$$
$$= 0 \quad z = 0, L, \tag{A3}$$

where A and B are constants. When $L/a \rightarrow 0$ and $\omega \rightarrow 0$, it can be analytically shown that $\omega \ll \omega_{mi}$ and that

$$A \cong -d_{31}\sigma/[h(1-\sigma)] \tag{A4}$$

and

$$B \cong d_{31}(a/h)/(1-\sigma). \tag{A5}$$

It can be easily verified by substituting Eqs. (A2)–(A6) into Eq. (8) that the static equations are satisfied. However, as observed earlier, the boundary conditions are satisfied only because of the discontinuity in $W(z,\omega)$.

Equations (A4) and (A5) can be obtained by first showing that

$$\Omega_{m1} \cong 1 - \sigma^2, \tag{A6a}$$

$$\Omega_{m2} \cong (k_m a)^2, \tag{A6b}$$

$$V_{m1} \cong \frac{4d_{31}a}{\pi h(1-\sigma)m}, \tag{A7a}$$

$$V_{m2} \cong 0, \tag{A7b}$$

$$V_{m1}C_{m1} \cong \frac{4d_{31}\sigma L}{\pi^2(1-\sigma)m^2}, \tag{A8a}$$

and

$$V_{m2}C_{m2} \cong 0, \tag{A8b}$$

when L/a and ω tend to zero. Then, the expression for A can be obtained by using Eq. (A8) and

$$1 + 1/9 + 1/25 + 1/49 + \cdots = \pi^2/8 \tag{A9a}$$

to evaluate the series expression for $U(z,\omega)$ in Eq. (A1) at $z=0$ and equating the result to the expression for $U(0,\omega)$ in Eq. (A2). Similarly, the expression for B can be obtained by using Eq. (A7) and

$$1 - 1/3 + 1/5 - 1/7 + \cdots = \pi/4 \tag{A9b}$$

to evaluate the series expression for $W(z,\omega)$ in Eq. (A1) at $z=L/2$, and equating the result to B.

[1] W. Soedel, *Vibrations of Shells and Plates* (Marcel Dekker, New York, 1993).

[2] P. R. Stepanishen, ''Modal coupling in the vibration of fluid-loaded cylindrical shells,'' J. Acoust. Soc. Am. **71**, 813–823 (1982).

[3] P. R. Stepanishen and D. D. Ebenezer, ''An eigenvector method to determine the transient response of cylindrical shells in a fluid with uniform axial flow,'' J. Acoust. Soc. Am. **89**, 565–574 (1991).

[4] J. F. Haskins and J. L. Walsh, ''Vibrations of ferroelectric cylindrical shells with transverse isotropy. I. Radially polarized case,'' J. Acoust. Soc. Am. **29**, 729–734 (1957).

[5] D. D. Ebenezer, ''Three port parameters and equivalent circuit of radially polarized piezoelectric ceramic cylinders of finite length,'' J. Acoust. Soc. Am. **99**, 2908–2912 (1996).

[6] R. Holland and E. P. EerNisse, ''Accurate measurements of coefficients in a ferroelectric ceramic,'' IEEE Trans. Sonics Ultrason. **SU-16**, 173–181 (1969).

[7] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Chap. 6.3.

[8] R. Holland and E. P. EerNisse, *Design of Resonant Piezoelectric Devices*, MIT Research Monograph No. 56 (MIT, Cambridge, 1969), pp. 99, 146, 153.

[9] D. D. Ebenezer and A. J. Sujatha, ''New methods to characterize radially polarized piezoelectric ceramic cylindrical shells of finite length,'' J. Acoust. Soc. Am. **102**, 1540–1548 (1997).

[10] O. G. McGee, A. W. Leissa, C. S. Huang, and J. W. Kim, ''Vibrations of circular plates with clamped V-notches or rigidly constrained radial cracks,'' J. Sound Vib. **181**, 185–201 (1995).

Smart acoustically active surfaces

F. Douglas Shields and L. Dwynn Lafleur[a]

National Center for Physical Acoustics, Coliseum Drive, University, Mississippi 38677

The use of such materials as PVDF and piezorubber in making smart acoustically active surfaces is discussed and demonstrated. By using one layer as a sensor and applying the signal detected in this layer through an amplifier and filter to a second driving layer it is possible to control either the reflection or transmission of plane waves striking the surface normally. By using two driving layers, both transmission and reflection can be simultaneously controlled. When transmissions are eliminated the control of reflection is independent of changes in backing impedance. By adding a second sensing layer along with two driving layers, reflections and radiation from a vibrating surface can be eliminated simultaneously. Stability problems involved in using pressure and/or velocity sensors in various control algorithms are discussed. © *1997 Acoustical Society of America.* [S0001-4966(97)00309-3]

PACS numbers: 43.38.Ar, 43.38.Fx, 43.40.Vn [PJR]

INTRODUCTION

The recent production of thin sheets of polymer and composites with piezoelectric properties has made possible the development of smart acoustically active surfaces. In this paper we discuss some of the problems involved in using these surfaces to control reflection, transmission, and radiation of plane waves.

The basic idea for such a surface is seen in Fig. 1(a), which diagrams plane waves striking a plane surface normally. The signal generated in a sound-sensing layer is amplified and filtered and applied to a driving layer. By controlling the complex gain in G, it is possible to adjust the impedance presented by the compound surface to the incident waves. For example, either reflections or transmissions can be eliminated.[1,2] By adding a second amplifier-filter, G_2, and a second driver as in Fig. 1(b), it is possible to simultaneously reduce both reflections and transmissions. By adding an additional sensor on the back surface, as in Fig. 1(c), it is possible to control simultaneously reflections and radiation.

If G_2 is adjusted to eliminate transmission, then the performance of the first layer in controlling reflections is independent of the backing impedance upon which the compound surface is mounted. (If there is no transmitted sound, the back surface is not vibrating. If it is not vibrating, the backing impedance can have no effect.)

The functioning of the second driver layer can be understood by considering the long wavelength limit where the motion of the second layer is considered to be the sum of its center of mass motion and the internal electrostrictive motion.[3] If these two sum to zero at the back surface of the second layer, its front surface will be moving with twice the amplitude of the center of mass. This is illustrated in the diagram in Fig. 2. The center of mass motion S_{CM} plus S_2 add to zero, but S_{CM} and S_1 add to $2S_{cm}$. This illustrates the problem of utilizing the second layer to make the operation of the first layer independent of backing impedance. While

the impedance seen by driver A is independent of the backing impedance, the constant impedance seen by A is approximately equal to one half that present by the mass of layer B with air backing. The difficulty in maintaining stability when controlling reflections in liquids from such a ''soft surface'' is discussed below.

I. DIFFICULTIES OF CONTROL

Two difficulties in making such smart surfaces are immediately obvious. The first is making the system respond fast enough and the second is feedback. Consider the timing problem first. For nonperiodic sound the driver must respond to the sound as it strikes its surface. If the control introduces appreciable delay, it will be necessary for the sensor to get an ''up stream'' look at the sound. If resonances are present in the system, the control amplifier filter G in Fig. 1 will need to have a complicated frequency response. The more resonances there are, the more complicated this transfer function is. A digital filter is the obvious answer. The filter, analog or
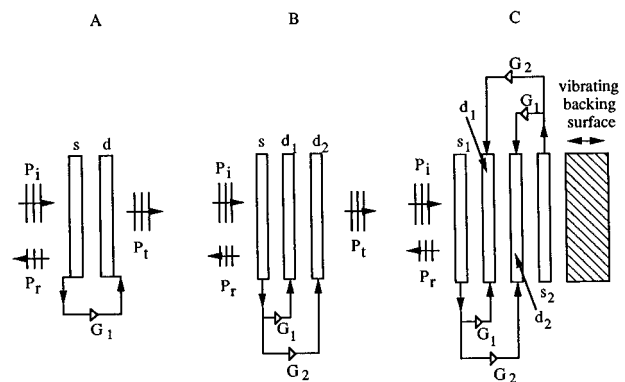
FIG. 1. (a) Smart acoustically active surface capable of controlling either reflection or transmission. (b) Smart acoustically active surface capable of simultaneously controlling both reflections and transmissions. When adjusted to eliminate transmission, the control of reflection will be independent of backing impedance. (c) Smart acoustically active surface capable of simultaneously controlling both reflection and radiation from a vibrating surface.
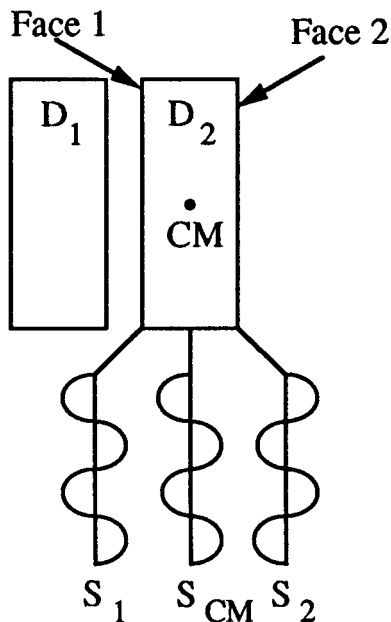
FIG. 2. Phase relations of the surface vibrations of a driving layer when cancelling sound transmission.

digital, will likely introduce time delays making broadband control difficult. The more complicated the transfer function, the more coefficients needed in a digital filter and the greater time delay introduced.

In order to discuss the second problem, that of feedback, it is necessary to consider different control algorithms.[4] In what follows consideration is limited to instabilities associated with the acoustic feedback path through the transducer layers and the medium in contact with the sensor. There are other paths (electromagnetic and structural) which complicate stability problems. It is possible to develop algorithms that make the sensor signal proportional to the total sound pressure or particle velocity at the sensor or to some combination of the two. A number of the possibilities are discussed below.

A. Sensor signal proportional to total pressure

The total sound pressure at the sensor, P, can be divided into incident pressure, P_i; reflect pressure when the driver is shorted, P_r^0; and pressure in a radiated wave generated by the driver, P_r^a. The latter is the filter transfer function, G, times whatever variable is used as the control signal. If that quantity is the total pressure, then the reflection coefficient, R, defined as the pressure amplitude in the wave leaving the surface to that arriving at the surface, is given by

$$R = \frac{P_r^0 + P_r^a}{P_i},$$

$$R = \frac{P_r^0 + G(P_i + P_r^0 + P_r^a)}{P_i}, \qquad (1)$$

$$R = \frac{R_0 + G}{1 - G},$$

with $R_0 = P_r^0/P_i$.

The reflection goes to zero when $G = -R_0$, but it goes to infinity at $G = 1$. The latter is obviously a condition of instability.

B. Sensor signal proportional to P_i

For plane waves, a signal proportional to the sound pressure in the incident wave can be obtained if the sensing layer combines a signal proportional to the total pressure with one proportional to ρc times the particle velocity of the surface. The total pressure, P, is $P_i + P_r$, but the total particle velocity for plane waves traveling in opposite directions is $u = u_i - u_r$. Therefore, since for plane waves $P = \rho c u$, $P + \rho c u = P_i + P_r + P_i - P_r = 2P_i$. In this case $P_r^a = 2GP_i$ in Eqs. (1) and

$$R = R_0 + 2G. \qquad (2)$$

In this case $R \rightarrow 0$ when $G \rightarrow -R_0/2$ and to infinity as G goes to infinity.

C. Sensor signal proportional to P_r

In this case $P_r^a = G(P - \rho c u) = C - (2P_r)$, and

$$R = R_0 + 2GR, \qquad R = \frac{R_0}{1 - 2G}; \qquad (3)$$

$R \rightarrow 0$ as $G \rightarrow \infty$ and $R \rightarrow \infty$ as $G \rightarrow \frac{1}{2}$.

D. Sensor signal proportional to $P_i + P_r^0$

Another control algorithm that is fairly easily implemented with a digital filter involves subtracting from the pressure sensor response that part that results from the driver. In this way it is possible to remove the basic plane wave feedback instability when using a simple pressure sensitive sensing layer. The procedure for doing this is to "model" the feedback path by putting out a known signal on the driver and digitizing the sensor response. With this model the digital filter can be programmed to subtract out the sensed pressure signal the part due to the driver. In the experiments described below this is done by programing the digital signal processor (DSP) to output a unit step in voltage to the driver. The digitized signal coming off of the sensor then becomes the needed filter coefficients for removing the driver signal. This procedure is explained by the following equation:

$$y(n+1) = G(x_n - \Sigma_i R_i x_{(n-i)}),$$

where $y(n+1)$ is the output from the DSP to the processor at time $n+1$, and x_n is the signal read into the DSP at time n. The R_i are the coefficients obtained by outputting a unit step from the DSP to the actuator and digitizing the response of the sensor. The response coefficients can either be obtained "off line" in the absence of an external signal or by averaging a low-level known signal over a period of time in the presence of uncorrelated sound.

By this procedure a signal is obtained that is proportional to $P_i + P_r^0$. Then

$$R = R_0 + G(1 + R_0). \qquad (4)$$
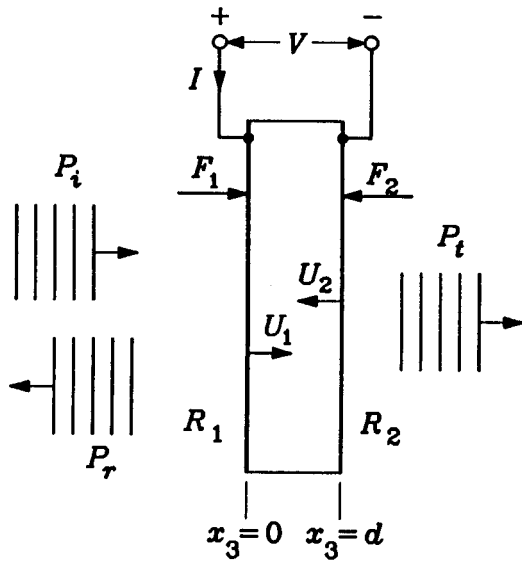
FIG. 3. Diagram of an active surface showing sign conventions for acoustic waves, surface velocities and forces, and the applied voltage.

Here, R goes to zero when $G = -R_0/(1+R_0)$ and increases to infinity as G goes to infinity.

(The functional dependence of R upon G for the various control algorithms discussed above has been demonstrated in the three-dimensional plots displayed in Figs. 4 and 5. These figures are explained in the next section.)

E. The effect of backing impedance upon stability

For the case of a single driver layer, the stability problem is greatly influenced by the backing impedance. This also is demonstrated in the figures in the next section. However, the general nature of the problem is understood by considering the two cases, first where the surface is hard (impedance greater than ρc of medium) and, second, the case where the surface is soft (the surface impedance is less than ρc of the medium). In the first case, if reflections are to be reduced, the driver drives the surface to make it softer. In this case the feedback to a pressure sensor is inverse, i.e., it decreases the sensed signal. To a velocity sensor, on the other hand, the feedback would be regenerative since the driver signal increases the sensed signal. For a soft surface, where the driver must make the surface harder to reduce reflections, inverse and regenerative roles for the pressure and velocity sensors are reversed.

II. CALCULATIONS FOR THE CASE OF A PIEZORUBBER TRANSDUCER

The above conclusions have been tested by theoretical calculations and, in some cases, experimental observations using piezorubber drivers and PVDF sensors. Reference 1 gives equations connecting the sound pressure in the incident, reflected, and transmitted waves to the voltage across and current to the driving layer for single-frequency plane waves striking the layer normally. The transducing layer is diagrammed in Fig. 3. The three equations connecting the five variables are

$$C_{11}P_i + C_{12}P_r + C_{13}P_t + C_{15}J = 0, \qquad (5)$$

$$-C_{22}P_i + C_{22}P_r + C_{23}P_t + C_{15}J = 0, \qquad (6)$$

$$C_{31}P_i - C_{31}P_r + C_{33}P_t + V + C_{35}J = 0. \qquad (7)$$

The values of the C's are given in terms of the mechanical impedances of the medium on each side of the transducer, the piezoelectric and the mechanical properties of the piezorubber, the thickness of the layer, and the frequency. For this ''three-port'' transducer, there are three equations and five variables, P_i, P_r, P_t, V, and J. Two of these can be assigned and the other three will be calculated as a function of R_1 and R_2, the impedances of the medium in contact with faces one and two, respectively. If a second driving layer is added, there results four equations and seven unknowns [see Eq. (5) of Ref. 1]. Equations (5)–(7) are used below to investigate the receiving sensitivity and the transmitting efficiency of a single layer of piezorubber. They are also used to show the stability characteristics of the different control algorithms discussed above.

A. Use of a single layer as a transducer

Here two special cases are considered. (1) The transducing layer acts as a receiver, for which we set $J = 0$, and (2) the transducer acts as a transmitter, for which we set $P_i = 0$. In each of these cases, two additional conditions will be considered: (a) medium two (on the transmission side of the transducer) is rigid, i.e., so $R_2 \to \infty$; and (b) medium two is free space (pressure release), i.e., $R_2 \to 0$ and $P_t \to 0$ with the ratio $P_t/R_2 = -u_2$ remaining finite, since u_2 is the complex amplitude of the velocity of the transducer surface bounding medium two.

Case 1a: Rigid backed receiver: Defining the receiving sensitivity as S as V/P_i, Eqs. (5)–(7) give for the case of rigid backing

$$S_R = -\frac{C_{31}}{C_{12}}(C_{11} + C_{12}),$$

where the C's are defined in the Appendix of Ref. 1. When the values of the C's are substituted in this equation this becomes

$$S_R = \frac{2(h_{33}/\omega)\sin\theta}{R_x \cos\theta + jR_1 \sin\theta}. \qquad (8)$$

Here, h_{33} is the piezoelectric stress constant, ω is the angular frequency, $\theta = kd$ where k is the wave number for plane waves in the transducing layer and d is the thickness of the transducing layer, $R_x = \rho_x V_x$ is the plane wave impedance of the transducing layer, and R_1 is ρc for the fluid in contact with side one. For the long-wavelength, low-frequency limit where the wavelength of the sound is much greater than the thickness ($kd \ll 1$), this equation reduces to

$$S_R \cong \frac{2h_{33}d}{R_x V_x}, \quad \lambda_x \gg d, \qquad (9)$$

where V_x is the plane wave velocity in the transducing layer. The receiving sensitivity is seen to be proportional to the thickness of the transducing layer and independent of frequency.

Case 1b: Free backed receiver: In this case

$$S_F = \frac{2h_{33}/\omega[1-\cos\theta]}{R_x\sin\theta-jR_1\cos\theta}, \tag{10}$$

which reduces to

$$S_F = j\frac{h_{33}d^2\omega}{V_x^2 R_1}, \quad \lambda_x \gg d. \tag{11}$$

In this case the sensitivity varies as d^2 and is also proportional to ω.

Case 2a: Rigid backed transmitter: In this case P_i is set equal to zero and Eqs. (5)–(7) are solved for P_r/J or P_r/V. They are labeled T_{JR} and T_{VR}, respectively. They give

$$T_{JR} = \frac{(h_{33}R_1/\omega)\sin\theta}{R_x\cos\theta+R_1\sin\theta}, \tag{12}$$

which reduces to

$$T_{JR} \cong h_{33}R_1d/V_xR_x \quad \text{for } \lambda_x\gg d. \tag{13}$$

For the sound pressure radiated per volt,

$$T_{VR} = \frac{h_{33}\varepsilon_{33}^s R_1\omega\sin\theta}{\omega d(R_1\sin\theta-jR_x\cos\theta)+jh_{33}^2\varepsilon_{33}^s\sin\theta}, \tag{14}$$

where ε_{33}^s is the clamped permittivity of the transducing layer. This reduces to

$$T_{VR} = j\frac{h_{33}\varepsilon_{33}^s R_1\omega}{R_xV_x-h_{33}^2\varepsilon_{33}^s}, \quad \lambda_x\gg d. \tag{15}$$

The long wavelength current transmitting efficiency is seen to be proportional to thickness and independent of frequency, while the voltage efficiency is proportional to frequency and independent of transducer thickness.

Case 2b: Free-backing transmitter: The free-backing transmitter efficiency with pressure release backing is

$$T_{JF} = \frac{h_{33}R_1(1-\cos\theta)}{\omega(R_x\sin\theta-jR_1\cos\theta)}. \tag{16}$$

For low frequencies this reduces to

$$T_{JF} \cong \frac{jh_{33}d^2\omega}{2V_x^2}, \quad \lambda_x\gg d. \tag{17}$$

For the free backing, the voltage transmitting efficiency is

$$T_{VF} = \frac{jh_{33}\varepsilon_{33}^s R_1R_x\omega\cos(1-\cos\theta)}{h_{33}^2\varepsilon_{33}^s R_x(1-\cos\theta)^2+G} \tag{18}$$

with

$$G = (h_{33}^2\varepsilon_{33}^s\sin\theta-\omega R_xd\cos\theta)$$
$$\times(jR_1\cos\theta-R_x\sin\theta),$$

which reduces to

$$T_{VF} = \frac{h_{33}\varepsilon_{33}^s R_xd\omega^2}{2V_x(h_{33}^2\varepsilon_{33}^s-R_xV_x)}, \quad \lambda_x\gg d, \tag{19}$$

from the equations

$$\frac{S_R}{S_F} = \frac{T_{JR}}{T_{JF}} = \frac{T_{VR}}{T_{VF}} = -j\frac{R_1}{R_x}\frac{\lambda_x}{d\pi}, \quad \lambda_x\gg d. \tag{20}$$

These equations demonstrate the well-known principle that transmitters and pressure sensors become insensitive when operating on soft backing.

B. Single-sensor–single-actuator smart surfaces

In this section Eqs. (5)–(7) will be used to demonstrate the stability characteristics of the different control algorithms discussed in the previous section. To calculate this dependence, V in Eq. (7) is set equal to a complex gain G times a sensitivity M times the control variable (P, $P-$feedback, $P+\rho cu$ or $P-\rho cu$). The complex reflection coefficient at the water transducer surface is calculated as a function of the complex quantity $G\cdot M$. For the case where the control signal is proportional to $P-$feedback, the equations are solved twice, once for the reflected pressure, P_r^0, when the transmitting layer is shorted, then V is set equal to GM times $P_i + P_r^0$. The calculations are made for transmitters composed of piezorubber 2 mm thick. The transmitting medium is water. The electromechanical properties of the piezorubber are given in Ref. 1. In that reference these values have been shown to adequately represent the piezorubber. The velocity is obtained from a pressure gradient sensor composed of two pressure sensors each with a sensitivity M separated by a 2-mm layer of ''ρc'' rubber (rubber with the same acoustical impedance as water). All pressure sensors are assumed to be thin, transparent, and to the have the same sensitivity. This
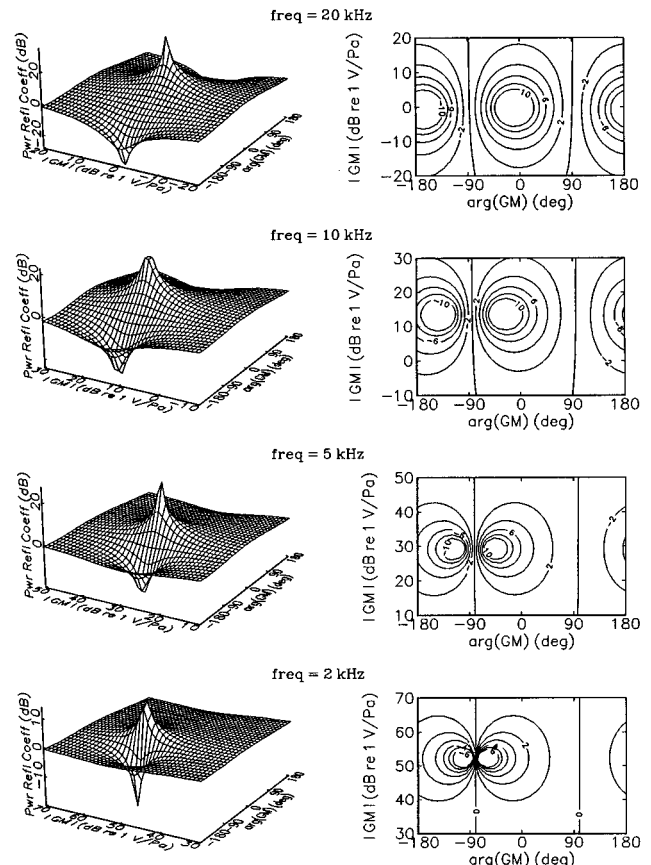
FIG. 4. Power reflection coefficient versus complex gain when the control signal is proportional to the total sound pressure on side 1.

means the sensitivity of the velocity sensor in the long wavelength limit will be proportional to the thickness of the ρc rubber separating the two pressure sensors. The results are plotted as a function of gain, G, times sensitivity, M, of the pressure sensor.

C. Total pressure control

Figure 4 shows the reflection coefficient versus gain when the input to G is proportional to the total pressure on side one. The results are shown in three-dimensional plots

(a) $\mathrm{P}+\rho\mathrm{cU}$ (b) $\mathrm{P-feedback}$ (c) $\mathrm{P}-\rho\mathrm{cU}$

freq = 20 kHz

freq = 10 kHz

freq = 5 kHz

freq = 2 kHz
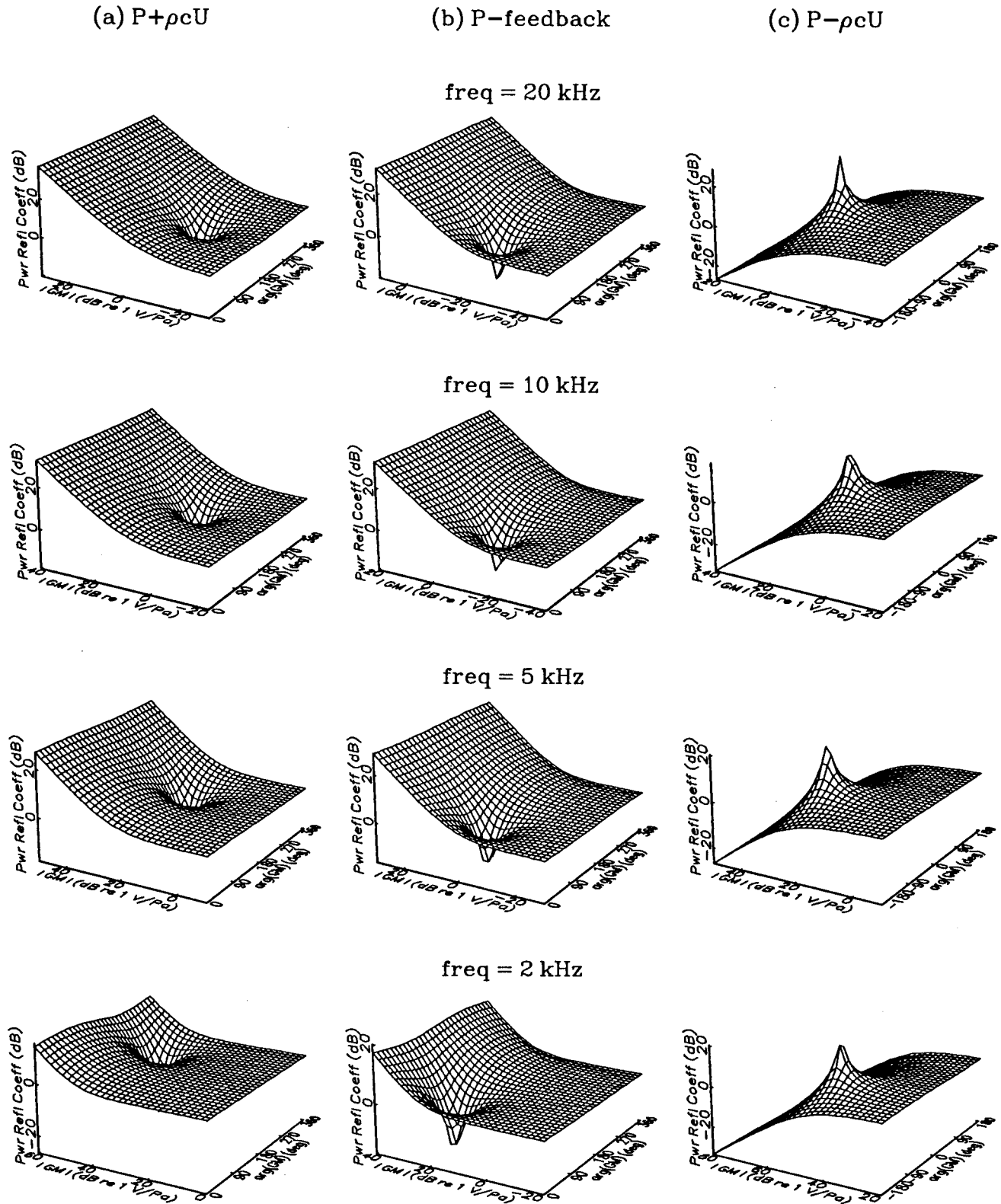
FIG. 5. Power reflection coefficient versus complex gain when the control signal is proportional to (a) $P + \rho c u$, (b) P-feedback, and (c) $P - \rho c u$.

(a) Power Reflection Coefficient

(b) Magnitude of P₂/Pᵢ
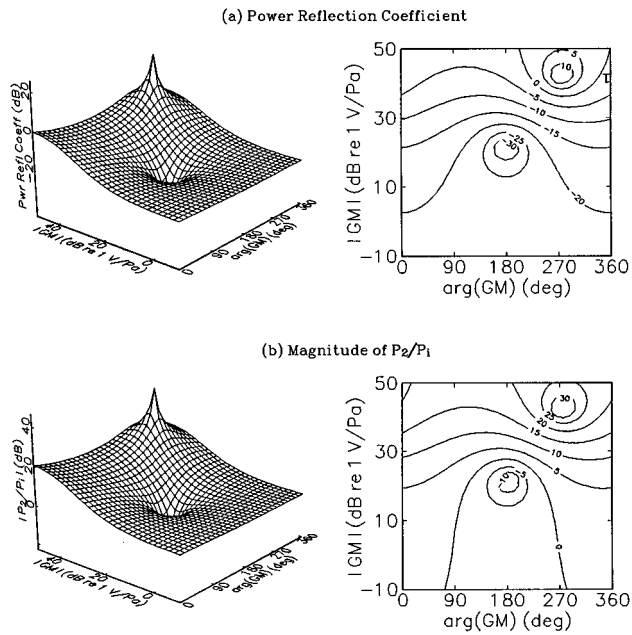
FIG. 6. Power reflection coefficient (a) and transmission coefficient (b) for a double layer of piezorubber. The frequency is 2 kHz. The figure illustrates the simultaneous cancellation of reflection and transmission.

(a) 2-cm Steel Backing

(b) Sea Water Backing

(c) 2-mm Steel Backing
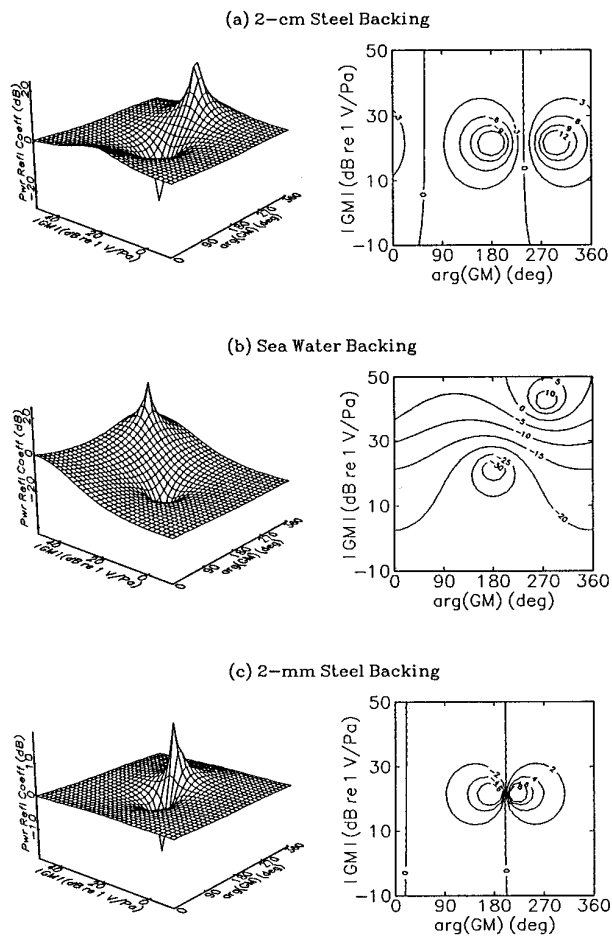
FIG. 7. Dependence of the reflection coefficient of the double layer upon backing impedance. The backing impedance decreases from *a* to *c*. While the position of instability moves in the complex gain plane, the position of the zero does not. The frequency is 2 kHz.

(a) $U_2=0$

(b) $U_2=-U_i$

(c) $U_2=jU_i$
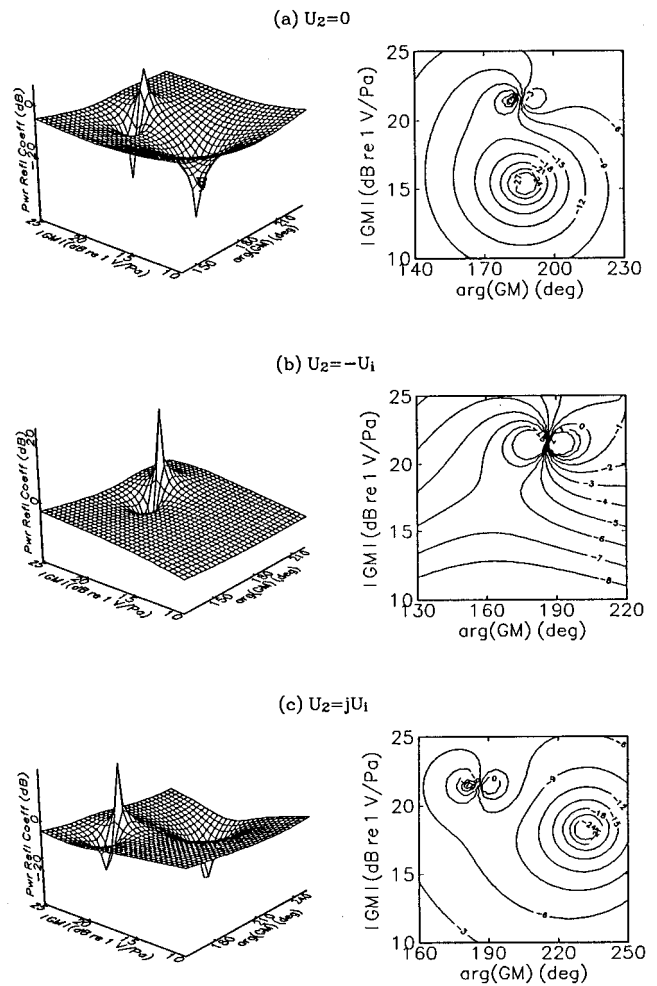
FIG. 8. Reflection and radiation from the active surface shown in Fig. 1(c) mounted upon a backing vibrating with a particle velocity, U_2. The figure demonstrates the possibility to simultaneously cancel reflection and radiation from a vibrating surface.
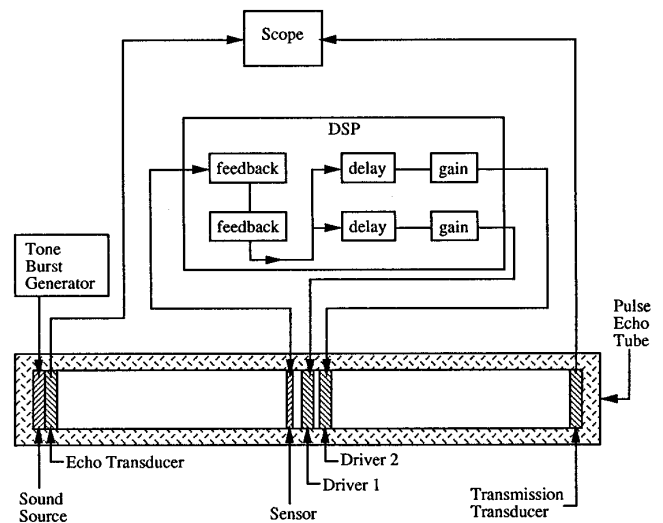
FIG. 9. Block diagram of the pulse echo tube and control circuit.
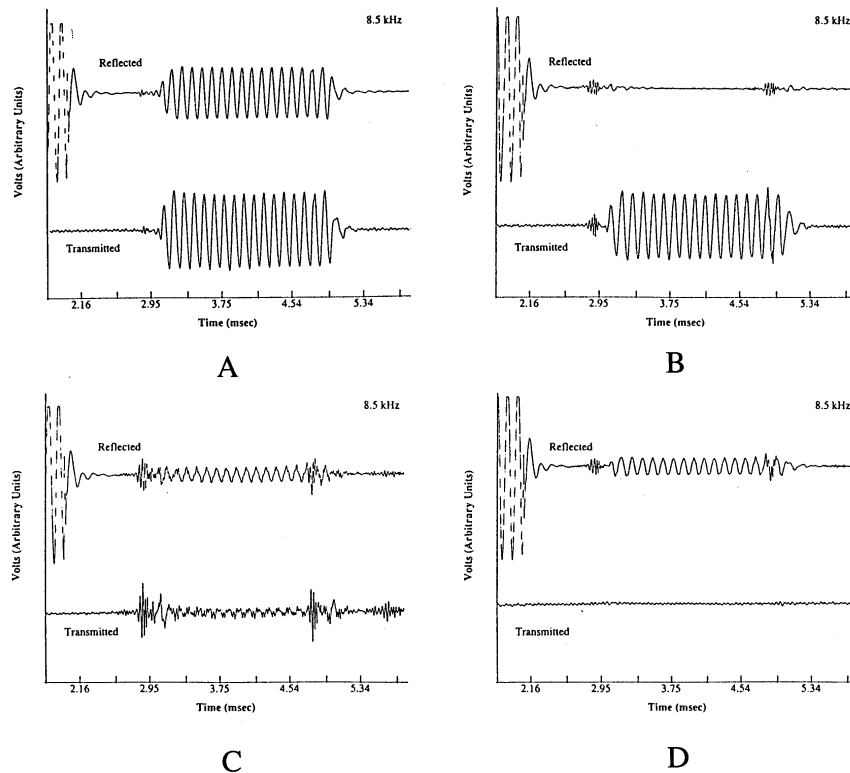
FIG. 10. Oscilloscope traces of tone bursts reflected from and transmitted through a smart bilayer diagrammed in Fig. 9. In figure A the bilayer is passive with oil on both sides. In B the DSP is programmed to eliminate only reflections. In C the DSP is programmed to simultaneously reduce transmission and reflections. In D 1 in. of styrofoam is added to the back of the smart bilayer. The styrofoam eliminates the transmitted burst completely, but has no effect upon the performance in reducing reflections. The figure shows the performance in reducing reflections is not dependent upon backing impedance if the bilayer is driven to eliminate transmission.

where the magnitude of the power reflection coefficient is plotted as a surface above a plane whose coordinates are the magnitude and phase of the complex gain. Then the same data is displayed as contour plots in two dimensions. The results are displayed for four frequencies 20, 10, 5, and 2 kHz. All of these frequencies are far below the piezorubber thickness resonance. The backing impedance is equal to that of 2 mm of steel backed by air. The poles and zeros and the general shape of the surface illustrate the problems involved in using this control algorithm discussed above. As the frequency decreases the backing becomes relatively softer and the pole and zero in the surface move close together, illustrating the problem of maintaining stability when the feedback becomes regenerative as it does when using a pressure sensor on a soft backing. Note also that much greater gains are needed for the same decrease in reflection as one goes from higher to lower frequencies.

D. $P+\rho c u$, $P-\rho c u$, and $P-$feedback control

Figure 5(a), (b), and (c) plots the reflection coefficient surfaces for the case where the input control signal is proportional to $P+\rho c u$, $P-$feedback, and $P-\rho c u$, respectively. The $P+\rho c u$ and $P-$feedback controls eliminate the acoustic instability (poles), but require careful control of the frequency dependence of the gain to accomplish much reduction in reflection. For the case of $P-\rho c u$, the careful timing is unnecessary but very large gains are required for significant reduction at low frequencies. Such large gains

present difficulties in controlling instabilities resulting from other types of feedback not considered here. Note in the $P-\rho c u$ graphs the magnitude of the gain is plotted in the reverse direction from the other plots.

E. Dual-driver and single-sensor smart surface

As discussed above, with two driving layers it is possible to simultaneously reduce both transmission and reflection. This is illustrated in Fig. 6 where the reflection and transmission coefficients at 2 kHz are plotted as a function of gain. This time Eq. (5) of Ref. 1 is used to calculate the gains. Both drivers are piezorubber 2 mm thick, and the dual driver has water on both sides. In the calculations $V_b=-(0.986973+j0.0941655)V_a$. This factor was calculated using Eq. (5) of Ref. 1 and solving for V_a and V_b when reflection and transmission are simultaneously eliminated. Here $V_a=GM\times P$, where P is the total sound pressure on side 1. The zeroes in the reflection and transmission coefficients then occur at the same value of gain, GM. The statement was made above that when this condition prevails, the operation of the surface is independent of the backing impedance. This is demonstrated in Fig. 7, where the reflection coefficient is plotted for three different backing impedances of increasing "softness." Notice that the zero in the reflection coefficient occurs at the same value of complex gain independent of backing impedance, but the pole and zero in the surface get closer together as the frequency goes down

and the backing impedance decreases. This is indicative of the increased difficulty in maintaining stability at low frequencies and soft backing.

F. Dual-transmitter dual-sensor smart surface

Consider now the compound surface diagrammed in Fig. 1C which consists of four layers, two sensors and two drivers. The signal from the first sensor to the two drivers is adjusted to cancel transmissions and reflections of waves penetrating the surface from the left. The signal from the second sensor is passed through an amplifier-filter with the same transfer function and applied to the two drivers for waves penetrating from the right. Figure 8(a), (b), and (c) shows that the same gain can produce zero reflection and radiation even when the backing surface is vibrating. The three figures show three cases at 2 kHz. One is when the backing surface is rigid ($U_2 = 0$). The second is when the backing surface has the same particle velocity as the incident wave, and the third is for the case where the backing surface has this velocity shifted in phase by 90°. The surfaces in Fig. 8(a) and (c) are seen to have two zeros and one pole. The second zero is not shown in Fig. 8(b) because the increase in scale of the horizontal plane does not allow its inclusion. At 2 kHz the second pole and zero are very close together and it would be difficult, therefore, to maintain stability in this region. It should be noted, however, that the zero close to the pole occurs at the same gain independent of the vibration of the backing. Therefore the gain needed to cancel simultaneously reflections and radiation is independent of the motion of the backing for a given frequency.

III. EXPERIMENTAL DEMONSTRATION

In Ref. 1, experiments were reported in which single- and double-layered active surfaces were driven from an external source so as to cancel reflections or transmissions (single layer) and both reflection and transmission simultaneously (double layer). Similar equipment has been used to produce ''smart'' surface cancelling. For this purpose, a sensor layer is added, and the signal from this sensing layer is passed through a digital filter and used to drive single- and double-layer transmitters to cancel reflections, and/or transmissions. As discussed above, Fig. 9 shows a block diagram of the pulse echo tube and the electronic control. The tone burst is generated by a piezorubber disc in one end of a 4.385-m-long pulse echo tube that is 6.31 cm i.d. and is filled with caster oil (see Ref. 1 for a complete description). The active surface under study is placed in the center of the tube and the transmitted and reflected pulses are detected with piezorubber transducers at each end of the tube. The scope traces shown in Fig. 10A, B, C, and D display the signals coming from these transducers. Two oscilloscope traces are displayed. The top trace is the signal from the echo transducer, and the bottom trace is the signal from the transmission transducer. In Fig. 10A the surface is passive, the digital filter connecting the sensor and the drivers is turned off. In Fig. 10B a signal is applied to one of the drivers and the digital filter adjusted to eliminate reflections. In Fig. 10C it is adjusted to simultaneously reduce transmission and reflection. Figure 10D shows that the performance in reducing reflection is unchanged by changing the backing impedance from oil to pressure release. To obtain the pressure-release backing a 2.54-cm-thick disc of styrofoam is placed behind the active surface. As expected, the styrofoam eliminates the transmitted pulse but has no observable effect on the reflected pulse.

In these experiments the control algorithm implemented in the DSP subtracts feedback from the sensed signal as described above. The control signal, therefore, is effectively $P_i + P_r^0$.

IV. CONCLUSION

Smart acoustically active plane surfaces can be effectively used to control both the reflection and transmission of normally incident plane waves. With two driven layers it is possible to simultaneously control both transmissions and reflections. When driven so as to eliminate transmissions, the performance of the double-layered surface in controlling reflections is independent of the backing impedance. By adding a second sensor it is possible to simultaneously control reflection and radiation from both sides of the active layer.

Feedback causes instabilities that introduce operational difficulties. These difficulties increase at lower frequencies where backing impedance decreases and needed gains increase. If control signals proportional to both sound pressure and particle velocity are available, they can be combined to reduce problems in controlling stability. By using digital control it is possible to model the feedback path and remove feedback from the sensor signal. This reduces feedback and introduces greater stability.

[1] L. D. Lafleur, F. D. Shields, and J. E. Hendrix, ''Acoustically active surfaces using piezorubber,'' J. Acoust. Soc. Am. **90**, 1230–1237 (1991).
[2] T. R. Howarth, V. K. Varadan, X. Bao, and V. V. Varadan, ''Piezocomposite coating for active underwater sound reduction,'' J. Acoust. Soc. Am. **91**, 823–831 (1992).
[3] D. M. Photiadis, J. A. Bucaro, and R. D. Corsaro, ''Double layer activator,'' J. Acoust. Soc. Am. **96**, 1613–1619 (1994).
[4] R. D. Corsaro and R. M. Young, ''Influence of backing compliance on transducer performance,'' J. Acoust. Soc. Am. **97**, 2849–2854 (1995).

Broad beamwidth ultrasonic transducers for tri-aural perception

Herbert Peremans[a)]

*Department of Artificial Intelligence, University of Edinburgh, Forrest Hill, 5, Edinburgh EH1 2Q1,
United Kingdom*

As airborne, correlation-based, ultrasonic sensors are becoming more widely used for robotic applications, it becomes increasingly important to have a good understanding of how the transducer's filtering properties affect the performance of the sensor system. Using well-known results from acoustics, a simple, yet accurate, filtermodel for the Polaroid transducer is described. In the course of this derivation the reason for the accuracy of the often used ''moving piston'' approximation is also clarified. This filtermodel is then used to analyze the characteristic ''peak-doubling'' from which the correlation based range sensor used in the tri-aural sensor suffers. The predicted errors are compared with actual measurements by the real sensor system and are found to be in good agreement. Finally, the use of broad beamwidth transducers to avoid these errors is proposed, and it is and argued that this approach has some distinct advantages compared to alternative solutions. © *1997 Acoustical Society of America.* [S0001-4966(97)02509-5]

PACS numbers: 43.38.Bs, 43.60.Qv, 43.58.Vb [SLE]

INTRODUCTION

In sensor based robotics,[1] the distinct advantages of ultrasonic range sensors when compared with visual sensors: low-cost, minimal processing, mechanical robustness, small size, etc., makes them a popular choice. Most of these ultrasonic sensors are simple time-of-flight sensors, their operation resembling closely that of the Polaroid sensor.[2]

However, due to the increased availability of low-cost computation power, correlation based techniques, previously too expensive both in computational and in hardware terms, are becoming more widely used in recent, more advanced, ultrasonic sensors.[3–5] This shift from relatively straightforward, threshold detector based, ultrasonic sensors to the more sophisticated correlation-based ones requires a heightened awareness of the filtering properties of the transducers. In particular, it will be shown that special attention should be paid to the angular-dependent filtering properties of ultrasonic transducers.

Because of the widespread use of the Polaroid[2] transducer in airborne ultrasonic systems, we will illustrate our argument using results derived for that particular transducer. In the course of this derivation we will also show, based on work by Anke,[6] why the directivity of the Polaroid transducer is well approximated by the ''moving piston'' hypothesis.

The analysis in this paper will indicate that this angular dependent filtering of the transducer may interfere with the correct operation of correlation-based range sensors, as it does with ours.[7] Based on these results, we advocate the use of broad beamwidth transducers for triangulation-based sensor systems, like the tri-aural sensor,[4] that use correlation-based range sensors as basic building blocks.

Section I briefly introduces the correlation-based range sensor as well as the tri-aural sensor consisting of three of such range sensors. In Secs. II and III a simple, but nevertheless quite accurate, filtermodel for the Polaroid transducer is discussed. The predictions from this filtermodel are then compared with actual measurements in Sec. IV. Finally, Sec. V contains a discussion of the consequences of this model for high-performance ultrasonic measurement systems.

I. TRI-AURAL PERCEPTION

In this section we will briefly describe both our correlation-based range sensor as well as the tri-aural sensor; more detailed accounts can be found in Ref. 7 and Ref. 4, respectively.

A. Range sensing

The range sensor starts each new measurement by transmitting a coded pulse. For simplicity reasons, most commercially available ultrasonic range systems choose a gated sinusoid to be the transmitted signal. However, as discussed in Ref. 7, the signal $x(t)$, shown in Fig. 1, possesses several advantages for ranging purposes. It is a 13-bit binary Barker code[8] implemented by phase modulating a 50-kHz carrier and representing each bit of the code by four periods of the carrier. This transmitted signal has a central frequency f_0 = 47 kHz and a centralized bandwidth B_c = 12 kHz.[9] Its amplitude spectrum $|X(f)|$ is shown in Fig. 7(a), allowing for efficient transmission of the signal when using the Polaroid transducer, as we will see below.

The signal received by the sensor is then processed with a matched filter. After quadrature demodulation of the output of this matched filter, Fig. 8(a), a peak detector determines the arrival time of the echo by locating the global maximum. If multiple objects are present in the field of view of the sensor, the peak detecting algorithm gets more complicated, since it has to distinguish between sidelobes; see Fig. 8(a), and true peaks.

[a)]Electronic mail: herbertp@dai.ed.ac.uk

FIG. 1. The transmitted signal $x(t)$.

FIG. 3. Comparison of the measured ($f_c = 54$ and $B_c = 9.9$ kHz) and the idealized ($f_c = 53$ and $B_c = 6.4$ kHz) transfer function $|H_{tr/rec}(f)|$.

B. The tri-aural sensor

The tri-aural sensor, as shown in Fig. 2(a), consists of three transducers lined up and spaced equally apart. The central transducer is used both as a transmitter and a receiver; the two peripheral ones are used solely as receivers. The three transducers are grouped into three transmitter/receiver pairs. Each transmitter/receiver pair forms a range sensor operating as outlined above. The measurement cycle of the tri-aural sensor is made up of the following steps. First, the central transducer emits the coded pulse. Then, at each of the three receivers, the precise arrival times of all the echoes are extracted from the raw transducer signals, Fig. 2(b), as explained in the previous section. Note that all three range sensors are sharing the same transmitter. This results in a list of arrival times associated with every receiver. Finally, these lists of arrival times are combined to find those that pertain to the same reflector, resulting in a list of triples, Fig. 2(c). A triple consists of the three arrival times, one from every receiver, of the echoes returned by a single object. The triples can then be used to determine the positions, both distance and bearing, of all the objects in the field of view by triangulation.

When using this sensor system on a robot moving around in the real world, some of the conditions for correct operation are not always fulfilled. For instance, some objects might not be ''seen'' by all three receivers. Since ultrasonic reflection is mostly specular, i.e., mirrorlike, the echoes received by each of the three receivers originate from slightly different points on the reflecting surface. Hence, it is con-

ceivable that, for certain object shapes, one of the receivers might not receive any echo at all, and thus no triple could be formed for that particular object. Or, as happens more often, an object might be too far to one side of the transmitter and hence not be seen by the receiver at the other side. Also, highly overlapping echoes may confuse the peak detection algorithm and have it detect false peaks or not detect the true peaks. Nevertheless, the correspondence algorithm that combines echoes from the three receivers is often able to cope with missing echoes as well as spurious echoes. A full account of how the correspondence algorithm uses the redundant information from the triples to achieve this can be found in Ref. 10, but note that the detection of an echo at two receivers is sufficient to locate the reflecting object, assuming both reflections originate from closeby points on the reflector.

An entirely different kind of problem occurs when the reflecting object is placed in certain, well-defined, positions in the field of view of the sensor system. In that case the signal received does not resemble the one shown in Fig. 8(a) after matched filtering, but instead looks like the one shown in Fig. 8(b). It is clear that having two peaks, both displaced with respect to where the true, single, peak should be, is going to result in erroneous position estimates. Below, we show that in the case of the Polaroid transducer a very simple model of the filter characteristics of that transducer can quite accurately predict this double peak phenomenon.

II. FILTER OPERATIONS PERFORMED BY THE POLAROID TRANSDUCER

The filter operations performed by the transducer can be separated into two independent filters. The first one models the dynamic effects involved in transforming the electrical signal applied to the transducer into the movement of its surface and back again. The second filter models the filtering performed by the transducer due to its spatially extended nature.

A. Transforming electric power into acoustic power

We will denote the input/output behavior of a perfectly aligned transmitter/receiver pair by its transfer function $H_{tr/rec}(f)$. In Fig. 3 we show this transfer function for the

FIG. 2. The tri-aural sensor.
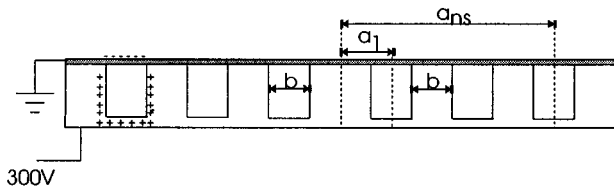
FIG. 4. An enlarged view of the multiple capacitors formed by the foil pressed against the grooved backplate.

Polaroid transducer; transmitter and receiver were separated by 1 m in the measurement.

For mathematical convenience we will approximate the measured transfer function with that of the filter, whose impulse response is specified by

$$h_{\text{tr/rec}} = \exp\left(-\frac{t^2}{2\sigma^2}\right)\sin(2\pi f_c t),\tag{1}$$

with f_c the frequency centroid and σ inversely proportional to the centralized bandwidth of the filter. Figure 3 shows both the measured transfer function and the idealized bandpass filter defined in Eq. (1). The parameters of the latter have been determined by a least-squares fit.

B. Angular dependent filtering by the Polaroid transducer

As long as the transmitter/receiver pair stays aligned, the behavior of the transducer is defined completely by the transfer function described above. If they are no longer aligned, additional filtering will take place because the transducer is not a point source.

The Polaroid transducer is basically a capacitor.[2] The gold plated side of its membrane serves as one plate of the capacitor, while the plastic side of the membrane is pressed against the other plate, the fixed backplate. In the backplate, concentric grooves are cut out at regular distances (Fig. 4). The membrane pressing against the grooves creates multiple, largely independent, air-filled capacitors, one for each ring-shaped groove. Other ways of constructing the transducer's backplate and how it influences the characteristics of the transducer are discussed by Rafiq.[11] The movement of the surface will be in accordance with this multiple capacitor model: concentric rings of moving membrane overlaying the grooves interlaced with stationary rings. As a first approximation, radius dependency of the velocity of the moving parts of the membrane is neglected.

Defining directivity as the ratio, expressed in decibels, of the pressure amplitudes for off-axis and on-axis observation points, Anke[6] shows that the movement of the membrane described above results in a directivity $D(\theta)$ given by

$$D(\theta) = 20\log\left(\frac{\left|\sum_{n=1}^{n=n_s} a_n J_0(ka_n \sin(\theta))\right|}{\left|\sum_{n=1}^{n=n_s} a_n\right|}\right),\tag{2}$$

with J_0 the zeroth order Bessel function and a_n, $n \in \{1,\cdots,n_s\}$, defined as shown in Fig. 4. We have measured these parameters on the backplate of a Polaroid transducer resulting in: $n_s = 34$ and $a_n = (n-1)0.5 + 0.375$ mm. The
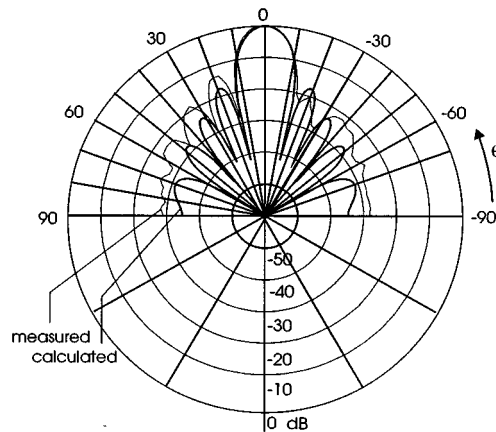
FIG. 5. Calculated and measured directivity of the Polaroid transducer at $f = 50$ kHz.

calculated directivity for the Polaroid transducer at a frequency of 50 kHz is shown in Fig. 5, together with the one we actually measured. The measured directivity agrees reasonably well with the predicted one, the asymmetric nature of the measurements most probably being due to inhomogeneities in the membrane and/or a nonuniform tension caused by asymmetries in the clamping of the membrane. We also notice that the zeros in the directivity pattern are slightly further apart than would be expected from the model. This can be explained by noting that the movement of the membrane, instead of being uniform, will taper off toward the edge where the membrane is clamped. This effectively reduces the size of the transducer and thus broadens the directivity pattern.

In most transducer models for electrostatic transducers the movement of the transducer surface is assumed to be uniform: the ''moving piston'' approximation. In the case of the Polaroid transducer, where the surface is formed by a conductive membrane, this assumption is not obviously valid. Hence, we have also compared the theoretical directivity with that of a moving piston:[12]

$$D_{\text{piston}}(\theta) = 20\log\left(\frac{|J_1(ka \sin(\theta))|}{|ka \sin(\theta)|}\right).\tag{3}$$

As can be seen from Fig. 6, they are virtually identical, showing that the ''moving piston'' approximation is indeed valid in the case of the Polaroid transducer.

Since the tri-aural sensor described in Sec. I, as well as most other correlation based systems, does not emit a pure tone, it is customary to consider the angle θ in Eq. (3) as a parameter and the frequency f as the independent variable. Hence, the spatially extended nature of the transducer can be modeled by the angular dependent filter

$$H_{\text{rad}}(f;\theta) = \frac{J_1(2\pi f a \sin(\theta)/v_s)}{2\pi f a \sin(\theta)/v_s}.\tag{4}$$

From Eq. (4) it is clear that the filter responses corresponding to different values of θ are all scaled versions of the same
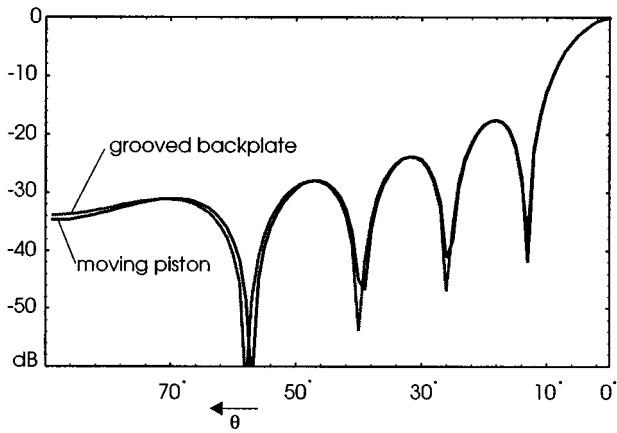
FIG. 6. Calculated directivity and directivity of a moving piston at $f = 50$ kHz.

FIG. 7. (a) The signal spectrum and the amplitudes of the transfer function for different viewing angles. (b) The received signal for $\theta = 2°$. (c) The received signal for $\theta = 14°$.

underlying transfer function. Examples of the effect of this transfer function for $\theta = 2°$ and $\theta = 14°$, i.e., the angle corresponding with the first zero in the 50-kHz directivity pattern, are shown in Fig. 7(a).

III. BEHAVIOR PREDICTED BY THE FILTER MODEL

In Sec. I we noted that the range sensor suffered from a characteristic peak doubling for specific angles. This faulty behavior of the sensor system can be readily explained by the filter model proposed above.

If we denote by $f_n(\theta)$ the frequency of a notch corresponding to a particular viewing angle θ, we can approximate Eq. (4) by linearizing it in the neighborhood of $f_n(\theta)$. This results in

$$H_{\rm rad}(f; \theta) \approx \frac{J_0(2\pi f_n(\theta)a \, \sin(\theta)/v_s)}{2\pi f_n(\theta)a \, \sin(\theta)/v_s} 2\pi(f - f_n(\theta)) \frac{a}{v_s}$$
$$\times \sin(\theta) = Cj2\pi(f - f_n(\theta)), \qquad (5)$$

where C includes all terms not depending on f. Next, denoting by $x(t)$ (central frequency f_0 and bandwidth B) the input to this notch filter and by $y(t)$ its output, we can write

$$x(t) = {\rm Re}[a_x(t)e^{j2\pi f_0 t}],$$
$$y(t) = {\rm Re}[a_y(t)e^{j2\pi f_0 t}], \qquad (6)$$

where $a_x(t)$, $a_y(t)$, the "complex analytic envelopes," [13] are not necessarily real-valued. Note that the bandpass filter $H_{\rm tr/rec}(f)$ as defined in Eq. (1) ensures that, irrespective of the transmit pulse used, $x(t)$ will indeed be a bandlimited signal. The Fourier transform of the analytic envelope is then given by

$$A_x(f) = \begin{cases} 2X(f + f_0), & |f| \leq B/2 \\ 0, & {\rm otherwise.} \end{cases}$$

Hence, we can rewrite the conventional filter equation

$$Y(f) = H_{\rm rad}(f; \theta)X(f)$$

in terms of the Fourier transforms of the analytic envelopes as

$$A_y(f) = H_{\rm rad}(f + f_0; \theta)A_x(f), \quad |f| \leq \frac{B}{2}.$$

For $f_0 = f_n(\theta)$, as is the case for Fig. 7(c), and using Eq. (5), this results in

$$A_y(f) \approx Cj2\pi f A_x(f), \quad |f| \leq \frac{B}{2},$$

or, equivalently,

(a)

(b)
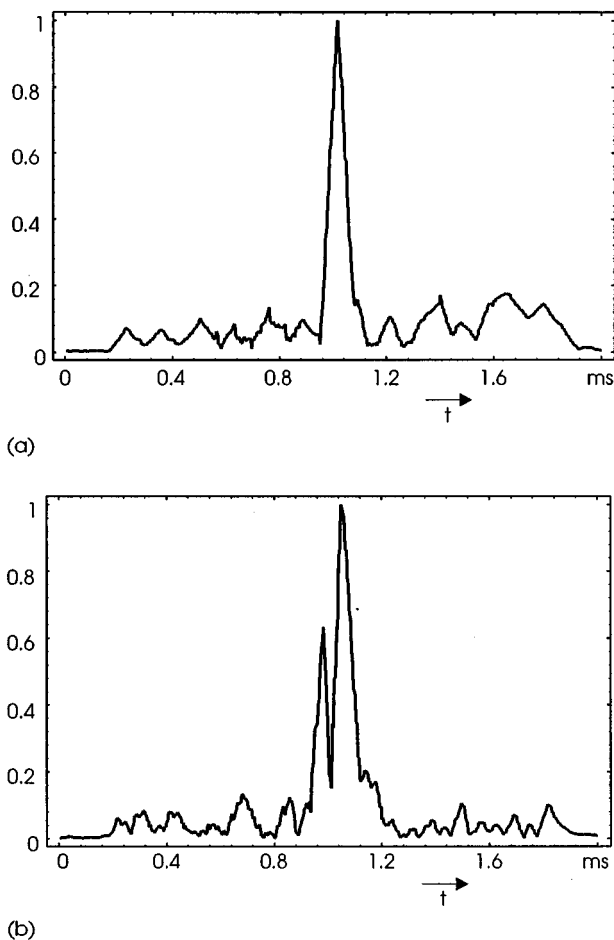
FIG. 8. Envelope of the measured output of the matched filter: (a) transmitter and receiver aligned, (b) transmitter and receiver displaced by 14°.

$$a_y(t) \approx C \frac{da_x(t)}{dt}.$$

Thus the notch filter introduced by $H_{rad}(f;\theta)$ for specific angles θ acts like a differentiator for the envelopes, explaining the appearance of a double peak for those angles.

IV. EXPERIMENTAL RESULTS

In Fig. 7 we display the received versions of the pulse transmitted by our ultrasonic range sensor for $\theta = 2°$ and $\theta = 14°$. For small angles, the first notch in the filter characteristic lies far above the frequency band used by the system. However, as the angle increases, the notches shift down toward lower frequencies and eventually enter the frequency band used by the sensor system, Fig. 7(a). From this moment on considerable filtering takes place, as can be seen from Fig. 7(c). Since our correlation based range sensor assumes the received signal to be a scaled and time-delayed version of the emission, it will not operate properly under those circumstances. This can be clearly seen from Fig. 8(b). Instead of a single peak located at the arrival time of the echo as in Fig. 8(a), we see two peaks, both displaced with respect to the true arrival time.

Furthermore, we know from Eq. (5) that the model predicts a similar peak doubling whenever a notch in the filter

(a)

(b)
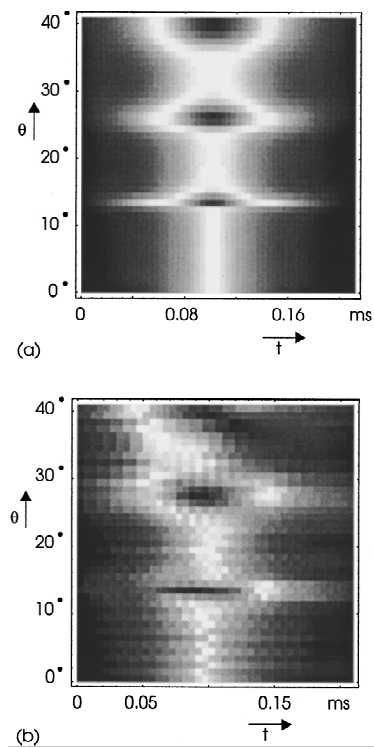
FIG. 9. A density plot of the central part of the envelope of the output of the matched filter after quadrature demodulation: (a) predicted by the model; (b) experimentally determined.

characteristic coincides with the central frequency of the transmitted pulse, i.e., the angle θ is close to one of the zeros in the directivity pattern at that frequency. To show that this is what actually happens, we have scanned the transmitter past a fixed receiver for viewing angles ranging from 0° to 41° in steps of 1° at a distance of 1 m, and compared the results with those predicted from the model. Both are summarized in Fig. 9, where we have shown the central part of the envelope of the output of the matched filter, for the different viewing angles. Note the characteristic doubling of the peak in the output of the matched filter whenever the viewing angle corresponds with a zero in the directivity pattern of the transducer. The first three zeros in the directivity pattern of the Polaroid transducer at 50 kHz are given by: $\theta = 14°$, 27°, and 41°. Both with regard to the number of zeros as well as to their respective positions, the model appears to correspond closely with the measurements. Of course, due to the idealizations introduced in the model, its output is smoother and more symmetric than the measured output. Nevertheless, the good correspondence between the predicted behavior of the sensor system and the actually observed behavior clearly shows the usefulness of the filtermodel.

V. DISCUSSION

From the analysis in the previous section we conclude that the faulty range estimates returned by the range sensor when the reflector is placed at specific positions can indeed be accurately predicted by a simple filtermodel for the ultrasonic transducer.

To minimize the influence of this filter, one can increase the bandwidth of the transmitted pulse and spread the energy over a wider frequency band as proposed in Ref. 5. Unfortunately, this is not a general solution, since for energy efficiency reasons the bandwidth of the transmission channel will be more or less constrained to that of the transducer itself. Therefore, when using a piezoelectric ceramic transducer, which has a narrow bandwidth, there is not much to be gained by increasing the bandwidth of the transmitted pulse.

Another approach keeps the transducer's transfer function constant by making sure that the reflecting object is always in a known position relative to the sensor: straight ahead, for example. There are indications that at least some biological sonar systems use this simple strategy for object recognition.[14] Once an object is detected in the periphery of the field of view, the bats, described in that paper, orient their heads towards the target, thus bringing it into the center of the beam. Only then is object recognition attempted. This strategy avoids erroneous results by actively controlling the measurement situation in such a way that the object of interest is within the region of correct operation and by ignoring the other, possibly incorrect, information by a focus-of-attention mechanism.

A third approach is the one proposed in this paper, i.e., to increase the range of bearings for which the system operates correctly by increasing the beamwidth of the transducer. The use of broad beamwidth transducers has the added advantage of resulting in a broader field of view for the sensor, meaning fewer measurements have to be gathered to inspect a region of interest. Because of the inherently slow measurement speed of ultrasonic sensors, this is an important advantage for mobile robot applications. However, the beamwidths should not be made too large, as having transducers with a smaller beamwidth allows the sensor system, by actively controlling the direction they are pointing in, to focus attention on certain reflectors without being confused by the echoes from other reflectors. This might explain why biological sonar systems have intermediate beamwidths; the 3 dB points of the directivity pattern of the bats described by Fuzessery et al.[15] are $\simeq 40°$ apart at 60 kHz. It could be objected that our approach requires that simple, one transducer sensor systems will have to be replaced by more complex, multiple transducers, ones that derive angular information from intensity ratios[16] or time differences,[4] for instance. But then it seems that this is the approach favored by nature when high-performance echo location is called for.

Alternatively, instead of trying to avoid the angular dependent filtering, one could consider it as a source of angular information. Indeed, the pinna, by increasing the directivity of the ear, is believed to contribute significantly to sound localization capabilities.[17] However, for robotic applications there are some distinct advantages to the approach we propose. In Ref. 5 it is noted that the angular dependent filter could in principle be used to determine angular information, but using time differences between right and left ear, as does the tri-aural sensor, is more accurate. Furthermore, by having to correlate the received signals with a different template for each angle of interest, the approaches described in Refs. 17 and 5 require far more processing than the tri-aural sensor using broad beamwidth transducers. Finally, to use the angular-dependent filtering to derive angular information, the sensor has to be calibrated for the particular transducers used. The tri-aural sensor, on the other hand, by using triangulation instead of a particular transducer property, is much less device dependent.

ACKNOWLEDGMENTS

[1] H. Everett, ''Survey of collision avoidance and ranging sensors for mobile robots,'' Robot. Auton. Syst. **5**, 5–67 (1989).

[2] C. Biber, S. Ellin, E. Shenk, and J. Stempeck, ''The Polaroid ultrasonic ranging system,'' in *67th Convention of the Audio Engineering Society* (New York, October 1980).

[3] C. Delepaut, L. Vandendorpe, and C. Eugène, ''Ultrasonic three-dimensional detection of obstacles for mobile robots,'' in Proceedings of the 8th International Conference on Industrial Robot Technology (Brussels), 483–490 (1986).

[4] H. Peremans, K. Audenaert, and J. V. Campenhout, ''A high-resolution sensor based on tri-aural perception,'' IEEE Trans. Robot. Autom. **9**, 36–48 (1993).

[5] L. Kleeman and R. Kuc, ''Mobile robot sonar for target localization and classification,'' Tech. Rep. Res. Rep. Ser. ISL-9301, Department of Electrical Engineering, Yale University, New Haven, Connecticut, September 1993.

[6] D. Anke, ''Luftschallwandler nach dem sell-prinzip fur frequenzen von 50 kHz bis 100 kHz,'' Acustica **30**, 30–39 (1974).

[7] K. Audenaert, H. Peremans, Y. Kawahara, and J. V. Campenhout, ''Accurate ranging of multiple objects using ultrasonic sensors,'' in *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*, Nice, May 1992, pp. 1733–1738.

[8] S. Golomb and R. Scholtz, ''Generalized barker sequences,'' IEEE Trans. Inf. Theory **11**, 533–537, October 1965.

[9] D. Menne and H. Hackbarth, ''Accuracy of distance measurement in the bat eptesicus fuscus: Theoretical aspects and computer simulations,'' J. Acoust. Soc. Am. **79**, 386–397 (1986).

[10] H. Peremans, ''A maximum likelihood algorithm for solving the correspondence problem in tri-aural perception,'' in Proc. of the IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems (Las Vegas), October 1994, pp. 485–492.

[11] M. Rafiq and C. Wykes, ''The performance of capacitive ultrasonic transducers using v-grooved backplates,'' Meas. Sci. Technol. **2**, 168–174 (1991).

[12] P. Morse and K. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).

[13] S. Rice, ''Envelopes of narrow-band signals,'' Proc. IEEE **70**, 692–699 (1982).

[14] D. Hartley and R. Suthers, ''The sound emission pattern and the acoustical role of the noseleaf in the echolocating bat, carollia perspicillata,'' J. Acoust. Soc. Am. **82**, 1892–1900 (1987).

[15] Z. Fuzessery, D. Hartley, and J. Wenstrup, ''Spatial processing within the mustached bat echolocation system: Possible mechanisms for optimization,'' J. Comp. Physiol. A **170**, 57–71 (1992).

[16] R. Kuc, ''Sensorimotor model of bat echolocation and prey capture,'' J. Acoust. Soc. Am. **96**, 1965–1978 (1994).

[17] P. Denbigh and P. Tollman, ''Beamforming by the cross-correlation analysis of received spectra,'' in *Adaptive Methods in Underwater Acoustics*, edited by H. Urban, 1985, pp. 439–446.

Sensor—actuator tile for underwater surface impedance control studies[a]

Robert D. Corsaro, Brian Houston, and Joseph A. Bucaro

Naval Research Laboratory, Code 7130, Washington, D.C. 20375-5350

An actuator—sensor tile has been developed for use in studies examining the issues involved in actively controlling the reflection and transmission characteristics of generic underwater structures. The tile is 25 cm square and contains a full area actuator, acoustic pressure sensor, and (acoustic particle) velocity sensor. This paper presents the predictive models used in the tile design. Issues addressed include transducer performance, spatial sampling, near-field sensing, internal resonances, and both direct and extraneous coupling mechanisms, all of which can contribute to complicate the system transfer functions. The acoustic characteristics of the final tile were evaluated at the NRL Laboratory for Structural Acoustics (LSA) and these results are compared with predictions. [S0001-4966(97)01409-4]

PACS numbers: 43.38.Hz, 43.40.At 43.30.Yj [SLE]

INTRODUCTION

Active control can be used to modify the acoustic scattering and radiation characteristics of structures. However for physically interesting structures, the acoustic field will typically be spatially complex. Accurate sensing and control of the dynamic acoustic field will then generally require a large number of acoustic sensors and actuators, a highly refined model of the plant (the physical system to be controlled), and a very fast controller (processor). For many acoustic applications, implementing this ''global'' approach would require a processor speed and complexity which is currently either unavailable or prohibitively expensive.

For structures satisfying certain requirements, an alternative simpler control approach can be considered. This approach partitions the surface of the structure into area elements and independently controls the response of each element. Such a ''local'' control approach is particularly advantageous if the algorithm to be implemented is identical for all surface area elements. The control strategy then reduces to applying multiple identical low-cost local surface area treatments.

We refer to these identical local surface area treatments as ''Smart Tiles'' by analogy to conventional ''passive tile'' (absorptive, decoupling, etc.) acoustic surface treatments. Each Smart Tile contains at least one sensor and at least one actuator. (A truly Smart Tile also contains an on-board controller; however this is beyond the scope of the present discussion.) The algorithm to be implemented (for any particular application) is required to be generic, such that tiles are structurally identical and spatially interchangeable.

This paper describes the development of a tile specifically designed for Smart Tile studies. This intended application involves examining the issues and capabilities related to using large arrays of these tiles to independently impose a local control algorithm to dynamically modify the acoustic fields present on generic complex structures in an underwater environment. Further discussion of these Smart Tile studies is beyond the scope of this paper, but some preliminary experimental results obtained for an array of 15 of these tiles on a complex structure have been discussed elsewhere.[1]

The present paper is limited to issues of tile design. Section I introduces the control approach as it relates to tile design requirements. Section II presents the predictive models used in transducer selection. Section III considers the key parameters related to acoustic coupling between transducers. Section IV compares predicted and measured acoustic characteristics of the resulting tile.

I. TILE FUNCTION

The tile must satisfy certain functional requirements related to its intended application in studies of active local control. As illustrated in Fig. 1, the control approach of interest partitions the surface of the structure into area elements called tiles, where each tile contains one acoustic projector or actuator, and at least one acoustic sensor. In its simplest implementation, the only information available to the controller comes from sensors located on the tile. Each tile also contains (or has an associated) controller or processor, which applies some control law (algorithm) in an attempt at enforcing a desired local condition.

With such limited sensor information available, the processor can only seek to impose a simple local acoustic boundary condition over the surface of the tile. In the specific case of interest here, the local control law we seek to impose is one which will control or dynamically select the reflection and transmission characteristics of the surface elements. Since these characteristics are determined by the surface acoustic impedance, the direct goal of this controller will be to actively force this impedance to a desired value. For example, one commonly sought condition is that of zero surface reflection at normal sound incidence. In a local sense this condition can be achieved if we force the normal acoustic input impedance of the surface to the value of the specific acoustic impedance of water, Z_w, which is simply the prod-
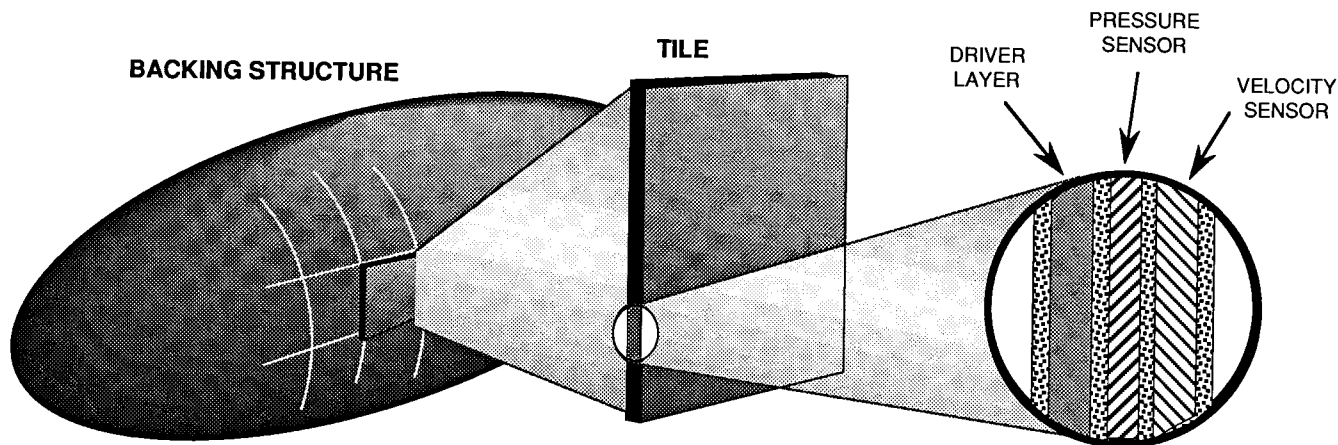
FIG. 1. Illustration of the control strategy.

uct of the density ρ and wave speed c in water. (This particular impedance condition is well suited to experimental studies since the degree to which it is achieved can be readily evaluated by simply measuring the normal reflection coefficient of the surface.)

The local control approach of interest is particularly suited to controlling the farfield reflected and/or transmitted acoustic energy. Hence the actuator associated with each tile should be shaded to generate a far-field actuator directivity function which matches that of the tile aperture. This directivity will generally be a function of frequency and test geometry (angle of incidence and observation, backing curvature, etc.). However it is observed that for normal (perpendicular) sound incidence (or radiation) on a (reasonably) flat surface, the directivity function of the surface region covered by the tile is identically that of a uniformly shaded actuator of the same size. Hence if properly controlled, for normal sound incidence a uniform actuator fully covering the tile surface can impose zero surface reflection at all observations angle.

For off-normal angles of incidence, the directivity function of a uniform actuator can be significantly different from the directional response of the surface. However, the importance of this mismatch can be largely tamed by limiting the tile dimensions. From a control viewpoint, the most problematic condition occurs near and beyond the first null in the scattering beam pattern where, if phase inversions are not properly included, serious controllability issues arise. Since these nulls occur when the projector dimension subtends multiples of one-half wavelength along the propagation wavefront, they can be avoided by restricting the (lateral) tile dimensions to a length no greater than one-half the wavelength of sound in water.

The tile designed here must be suitable for generic application on a wide variety of backing structures. If a controller is to actively force the value of the acoustic input (or output) impedance of the surface, the sensors selected must monitor this impedance to provide the relevant input signals to the controller. Since this impedance is the complex ratio of the acoustic pressure to velocity at the surface of interest, it follows that both acoustic pressure p and particle velocity v sensors are needed. For example, from simple consider-

ations, it has been shown[2] that to force a nonreflective condition, the task of the controller is to impose the condition

$$|p + \rho c v_n| \Rightarrow 0 \tag{1}$$

at all locations on the tile surface, where throughout this paper we will use the radiation convention in which positive velocity values indicate propagation away from the surface. Rigorously imposing this condition at all locations on the tile would require a complicated actuator shading function. However since we have already specified that the tile size is small (less than one-half wavelength), the variations in pressure or particle velocity over the aperture of the tile will not be large. Then it is sufficient for the pressure and velocity sensors to output an area averaged value over the tile surface.

II. TILE DESIGN

To obtain an integrated tile with well behaved and predictable performance characteristics, each of the three transducer layers must also be well behaved and understood. In this section we discuss each transducer component, and predict their acoustic characteristics using the center-of-mass model of Photiadis[3] in the notation of Corsaro.[4] Some additional information and data have appeared in a previously published summary[5] and is not repeated here.

A. Actuator

The actuator material selected for this tile is the 1-3 composite available from Material Systems Inc.[6] (MSI). This type of composite consists of an array of uniformly spaced piezoelectric ceramic rods, where the space between the rods is filled with a compliant material. Cover plates (printed circuit board material) are typically applied to the top and bottom surfaces to form the electrical connections and to provide an environmental seal.

The principal advantage of a 1-3 composite is its high acoustic output per applied volt. In a conventional disk-shaped transducer element, upon the application of a positive voltage, the thickness of the element increases (a positive d_{33} constant) but its radius correspondingly shrinks (a negative d_{31} constant). When such conventional elements are

used in projectors, unless other precautions are taken, the radial shrinkage cancels much of the effect of the thickness change and the resulting volume change or hydrostatic piezoelectric charge constant d_h is quite small—typically $10\%-15\%$ of the d_{33} constant for the familiar lead zirconate-titanate (PZT) type materials. The 1-3 composite approach removes the detrimental influence of this radial shrinkage by decoupling this radial motion from the surrounding fluid. This results in a value of the effective volume piezoelectric charge constant which can approach the d_{33} value of the material.

For the present study we used first-production 1-3 composite material which was formed by a (then) new injection-molding process developed by MSI under Office of Naval Research (ONR) contract. The piezoelectric ceramic used is PZT-5H™ with a d_{33} piezoelectric constant of 600×10^{-12} C/N. The piezoelectric ceramic rods are each 6.3 mm in length and uniformly spaced on a 2.5-mm grid, where the space between the rods is filled with a voided compliant polymer.

In estimating the expected far-field performance of this actuator using the center-of-mass model, we find it first convenient to define a function we will refer to as driver efficiency **D** which is an equivalent local voltage-to-pressure transfer function:

$$\mathbf{D} \equiv \frac{\rho c v_d}{E_d} = i\omega\rho c \zeta_d d_{\text{eff}}, \tag{2}$$

where v_d is the driven normal velocity of the front-face of the actuator, E_d is the applied drive voltage, ω and λ are the angular frequency and wavelength, A is the surface area of the front-face of the actuator, d_{eff} is the effective piezoelectric constant of the transducer material, and ζ_d is the fraction of the total dynamic thickness change which appears as a displacement of the front-face (for simple water immersion ζ_d is 0.5). The transmit voltage response (TVR) is then given as:

$$\text{TVR} = 20 \log\left\{ \frac{A}{\lambda} \mathbf{D} \right\}, \tag{3}$$

where by the usual definition[7] TVR is the far-field pressure produced by 1 V, projected back to a reference distance of 1 m using the spherical spreading law. The resulting predicted TVR for this actuator using a value of the effective piezoelectric constant of 400×10^{-12} C/N (discussed latter) is shown in Fig. 2.

B. Pressure sensor

The pressure sensor (hydrophone layer) uses a Piezorubber sheet (PR-307™ from NTK Corp. of Japan) as the transducer material. The predicted voltage-pressure transfer function, also called the pressure sensitivity, is defined as

$$\mathbf{S}_p \equiv \left(\frac{E_p}{p} \right) = g_h t_s \gamma_p, \tag{4}$$

where E_p is the sensor output voltage, t_s is the transducer thickness, and g_h is the hydrostatic piezoelectric voltage constant. The parameter γ_p is approximately unity (within 0.05
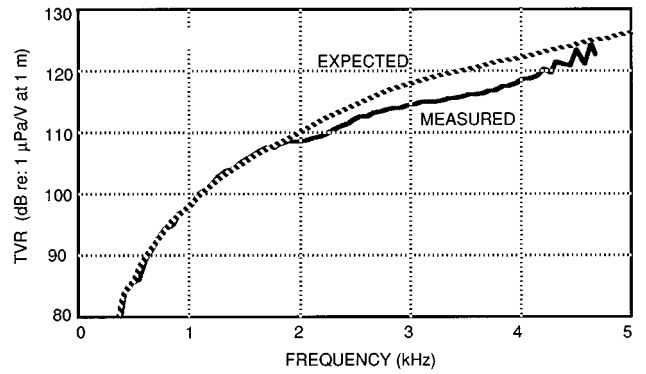
FIG. 2. Actuator free-field TVR of the assembled tile.

dB) for the sensor dimensions, materials, and bandwidth considered here.[4] Using the value for g_h given by the manufacturer (0.111 V m/N) and a typical thickness of 2.8 mm, the predicted sensitivity is -190 dB *re*: 1 V/μPa.

A key sensor parameter is the minimum detectable pressure change, or the noise floor of the sensor-preamplifier combination. The pressure sensor, including noise generators, can be represented using an equivalent network modeled after that of Tims[8] as illustrated in Fig. 3. The subscripts used are s, x, and n corresponding to the sensor, stray external contributions (cable, etc.), and preamplifier (equivalent input), respectively.

The signal output corresponding to an applied pressure is represented as a current generator I_p in parallel with a capacitor C_s and a resistor R_s. This signal I_p is not part of the noise contribution. The other three current generators shown produce the equivalent Johnson's noise current associated with their corresponding resistor:

$$I^2 = \frac{4k_B T_K}{R}, \tag{5}$$

where k_B is Boltzmann's constant (1.38×10^{-23} J/°K) and T_K is the absolute temperature. The electrical impedance of this circuit is the reciprocal sum (parallel combination) of the impedance of the total resistance R_T and the total capacitance C_T. The magnitude of this impedance is then

$$|Z_E| = \frac{R_T}{\sqrt{1 + \omega^2 R_T^2 C_T^2}}, \tag{6}$$

where C_T is the sum of the three capacitors, and the reciprocal of R_T is the usual reciprocal sum of the three resistors. The noise spectrum level at the output e_{out} (V/$\sqrt{\text{Hz}}$) is then given by

$$e_{\text{out}}^2 = e_n^2 + I_T^2 Z_E^2, \tag{7}$$
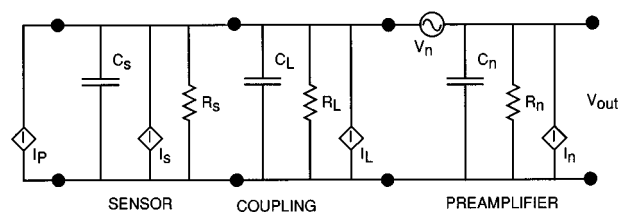
FIG. 3. Equivalent network model of the sensor.

where I_T^2 is the sum of the squares of the three noise currents. The total rms output noise level is then e_{out} multiplied by the square root of the frequency bandwidth Δf_{bw}. This noise floor for the hydrophone in terms of a minimum detectable pressure level is

$$P_{min} = \frac{e_{out}}{S_p} \Delta f_{bw}. \tag{8}$$

From such an analysis, we find that the noise level for this sensor is largely dominated by the input noise of the preamplifier selected. Using values for our Ithaco Model 1201 preamplifier,[9] the background noise spectrum level of the sensor is on the order of 35 dB *re*: 1μPa/$\sqrt{\text{Hz}}$ at frequencies above 1 kHz, which is below the acoustic noise floor of most test facilities. Using the sensor broadband with a 20-kHz operational bandwidth, the minimum detectable acoustic pressure is then $P_{min} = 76.3$ dB *re*: 1μPa.

C. Velocimeter

The area average velocity sensor was formed using an array of commercially available accelerometers. The output signals of these accelerometers are summed and integrated (using an on-board module) to obtain an output signal proportional to velocity.

While a dense accelerometer array is desirable, the number of accelerometers available is limited by cost. For this tile we elected to use an array of only four accelerometers. These are located on the tile diagonals with a (diagonal) separation distance of 20 cm, such that an array of tiles would be evenly sampled. This separation nominally provides at least $\lambda/2$ sampling (two accelerometers for each acoustic wavelength) for freely propagating waves at frequencies to 3.5 kHz for all incident or radiated angles.

For many studies and geometries, such as those using external insonifying signals and large-area actuators, the excitation of flexural waves is not a key issue. In studies where flexural waves are important, such as when the structure is excited by an internal noise source, the flexural wave sampling density limits the maximum useful frequency.

As an example, one backing structure available in our laboratory consists essentially of an air-backed 1.27-cm-thick steel plate. With the requirement of $\lambda/2$ sampling for flexural waves in the backing structure, from standard texts[10] we find the maximum useful frequency is given by:

$$f_{max} = \frac{\pi}{4\sqrt{3}} \frac{h_p c_p}{L_a^2}, \tag{9}$$

where c_p is the plate speed (5489 m/s for steel), h_p is the backing structure plate thickness, and L_a is the accelerometer separation. The $\lambda/2$ sampling frequency of the tile on this thickness of air-backed steel plate is 1.4 kHz. This would be adequate for most of the intended studies. Note however that since the actual stiffness of the laboratory backing structure is at least twice as great as this due to the inclusion of ribs and supports, we expect to obtain adequate sampling for internal-driven flexural waves to approximately 2 kHz.

Wilcoxon Model 759 accelerometers were selected for use, where these accelerometers were designed by Wilcoxon to meet Naval Research Laboratory (NRL) requirements of this application. Their main attributes[11] are high resonance frequency (25 kHz) and very low noise floor (nominally 0.10 μg/$\sqrt{\text{Hz}}$). They also have low sensitivity to off-normal (transverse) acceleration, and are waterproof. The sensitivity of these accelerometers is $S_a = 0.10$ V/g. The corresponding velocity sensitivity S_v for each accelerometer is then

$$S_v \left(\frac{\text{V}}{\text{m/s}} \right) = \frac{i\omega}{9.8} S_a. \tag{10}$$

The noise spectrum can be obtained from a fit to the manufactures specifications:

$$N_u(\mu\text{g}) = \sqrt{n_1^2 + \left(\frac{n_2}{f} \right)^2}, \tag{11}$$

with $n_1 = 0.10 \, \mu$g/$\sqrt{\text{Hz}}$, and $n_2 = 20 \, \mu$g $\sqrt{\text{Hz}}$, where the second term is significant only at frequencies below 1 kHz. Hence the noise spectrum level at 1 kHz (for example) expressed as an equivalent free-field pressure level is approximately 48 dB *re*: 1 μPa. This is on the order of the background acoustic level in typical laboratory evaluation tanks.

The minimum detectable velocity (noise floor) is the root mean square (rms) of the noise spectrum over the operating bandwidth (0.4–20 kHz). The corresponding noise figure is found to be nominally 15 μg rms (or an electrical output of 1.5 μV rms). The free-field pressure level corresponding to this minimum detectable velocity is then 91 dB *re*: 1 μPa@1 kHz, where due to the frequency dependence of the sensitivity, this minimum detectable equivalent pressure level is inversely proportional to frequency.

A further noise floor reduction of 6 dB results when the outputs of the four accelerometers are averaged. This is performed by an on-board electronic module (labeled ''E-VASE'') designed by NRL[12] to condition and sum the outputs of the four accelerometers. This module includes an integrator, such that the electrical output is linearly proportional to average velocity. The module contains eight amplifiers, and was constructed using surface mounted electronic components. It is packaged in a watertight nickel box measuring $3.8 \times 3.8 \times 1.8$ cm^3, which is then filled with silicone rubber (RTV-615™ from Dow Corning Co., Inc.). Of particular importance to control applications, the phase uniformity of this module is $\pm 5°$ over the frequency band 400 Hz–8 kHz.

III. TRANSDUCER COUPLING

The electrical coupling between the actuator layer and a sensor via acoustic mechanisms is defined in general terms as:

$$C_j \equiv \frac{E_{sensor}}{E_{drive}} = \alpha_i D_j S_j, \tag{12}$$

where α is defined here as the mechanical or acoustic coupling between actuator and sensor.

This electrical coupling is of particular importance in the design of a control system since it nominally corresponds to the inverse of the required neutralization gain. Hence smoothly varying coupling functions with high magnitudes are desirable, since they suggest that a stable low-gain control system can be configured.

A. Pressure sensor—actuator coupling

For the pressure sensor, the mechanical coupling α_p is unity if the actuator and sensor are very large. For a finite area radiator in close proximity to the sensor, it can be shown that the value of α_p is approximately that of the normalized radiation impedance of the actuator Z_{rad}.

A more nearly accurate value for α_p can be obtained by computing the pressure generated by the actuator (normalized by the forcing function $\rho c v_d$) at all locations in the plane of the pressure sensor, and then integrating this field over the area of the sensor. A computationally efficient way of computing this function is by approximating the actuator and sensor as circular, such that one can use the variable transform approach Archer-Hall and Gree[13] specifically as applied by Yen.[14] The approach finds that the (normalized) pressure at any field point from a disk radiator can be expressed by the single integral:

$$P_I = \left\{ \begin{bmatrix} 0, & b>a \\ 0.5, & b=a \\ 1, & b<a \end{bmatrix} e^{-ikz} \right.$$
$$\left. + \frac{1}{\pi} \int_0^\pi e^{-iks} \left(\frac{ab\cos(\phi) - a^2}{a^2 + b^2 - 2ab\cos(\phi)} \right) d\phi \right\}, \quad (13)$$

where the field point s is an axial distance z from the center of the disk radiator and distance b from the disk axis,

$$s \equiv \sqrt{a^2 + b^2 + z^2 - 2ab\cos(\phi)} \quad (14)$$

and where a is the radius of a disk (in this case) having an area equivalent to that of the (square) actuator.

B. Velocity sensor—actuator coupling

For the velocity sensor, the mechanical coupling α_v is also unity if the actuator and sensor are large and closely spaced. A more nearly accurate calculation can be based on the expression for the known pressure field at any location P_I as presented in Eq. (13). Then we can use Euler's relation to solve for the axial and radial components of the velocity vector at the accelerometer location. As illustrated elsewhere,[5] at our highest intended operating frequency of 4 kHz the magnitude of the axial velocity falls off significantly as one moves toward the edge of the radiator. However at the accelerometer location the measured axial velocity component is within 1 dB of the tile-area averaged response. At this location the radial velocity component at this frequency is substantial, and the overall velocity vector is pointing at an angle of 35° from normal. This imposes a requirement that the accelerometer selected have a low sensitivity to lateral accelerations.

C. Additional mechanical components

As illustrated in Fig. 4, in addition to the usual electrical, waterproofing, and bonding layers, the final tile construction includes two noteworthy construction features to smooth and simplify the transducer coupling transfer functions.

A transverse or lateral decoupler layer was included between the pressure sensor and the actuator to reduce extraneous direct coupling from lateral motions. Both of the transducer materials, NTK Piezorubber and the 1-3 composite, are expected to have transverse piezoelectric constants which are on the order of 10% of the longitudinal value. Hence if these two transducer elements were bonded together the ratio of the lateral to transverse coupling should be approximately −40 dB. This is not quite sufficient for some control investigations and applications. Hence we further reduced direct transverse coupling at the sensors by including the lateral decoupler. This decoupler consists of a thin nickel plate placed on a 3.2-mm-thick urethane rubber layer. From a simple finite element model,[15] it can be shown that this decoupler suppresses direct lateral motion by at least an additional 20 dB without substantially affecting longitudinal waves in the frequency band of interest.

An aluminum disk is also needed beneath each accelerometer to increase the stiffness of the underlying rubber composite structure and to increase the rotational stiffness of the accelerometer in the velocity field. Without the disk, the accelerometer sees an underlying stiffness which is dominated by the Young's modulus of the NTK Piezorubber and the rubber layer (encapsulant) covering the driver. The mass of the accelerometer will then give rise to a classical spring-mass resonance. Although strongly damped in water, this resonance is troublesome since in the present design it introduces a phase anomaly in the operating band of interest. The primary action of the disk is to increase the footprint of the accelerometer; this increases the frequency of this resonance moving it out of band. The maximum dimension of this disk is limited by the presence of flexural waves in the aluminum.

IV. EXPERIMENTAL RESULTS

A. Benchtop tests

The capacitance of this actuator was measured as 50 nF±3% with (2.0±0.2)% dissipation at 1 kHz. These values for a fully assembled actuator layer in a final tile are slightly greater than expected (43 nF), presumably due to the additional capacitance and resistance of the electrical connectors and shields.

An issue of particular importance to the tile design is its dynamic surface uniformity. Nonuniformity in the surface motion of the actuator will generate extraneous local mechanical interference which will complicate the actuator—sensor coupling and hence degrade performance in the intended Smart Tile application. Since the 1-3 composite actuator supplied by MSI was (at the time) a new transducer technology, some fabrication issues influencing uniformity had to be resolved prior to tile construction. For the final tile, measurements by NRL using Laser Vibrometry over a 1024 by 1024 grid determined that the variabilities in displace-
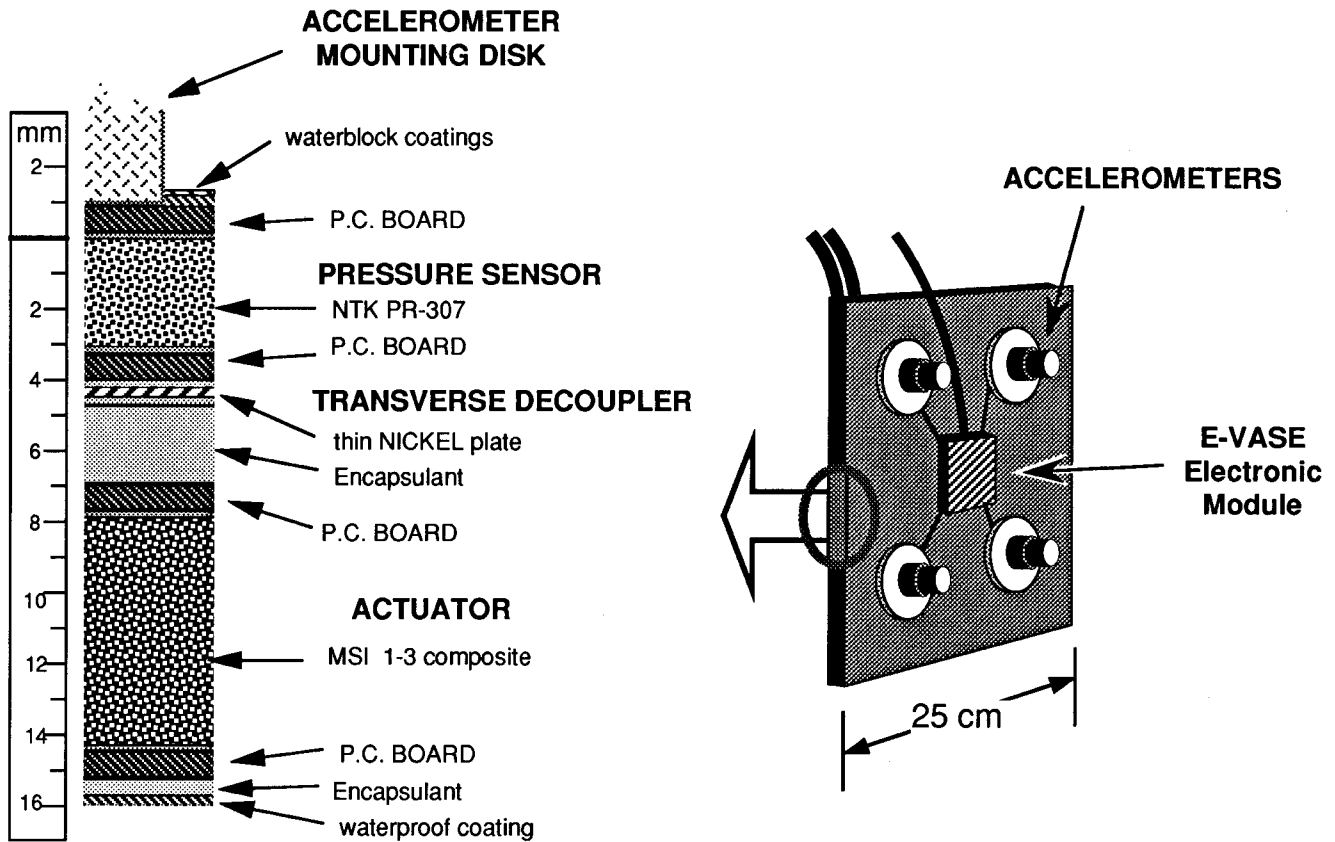
FIG. 4. Major components of the final tile design.

ment across the surface of the composite actuators was less than 1 dB.

The capacitance of the pressure sensor was measured as 10.9 ± 0.5 nF with $(2.7 \pm 0.2)\%$ dissipation. These are in agreement with values listed in the manufacturer's literature for the material alone (8.2 nF and nominally 4% dissipation).

The accelerometers were all supplied with calibrations. Spot checks using benchtop measurements verified these calibrations. Additional bench tests of the on-board E-VASE electronic module were performed with the hard-wired accelerometers during the development and certification of this module design.

B. Individual transducer performance

A series of measurements were performed on a fully assembled tile under simple water immersion free-field conditions. They were conducted at the Laboratory for Structural Acoustics. This is a highly instrumented acoustic pool facility at the Naval Research Laboratory (Washington, D.C.). The main element of the facility is a vibration-isolated temperature-controlled one million gallon fresh water tank that is instrumented with large-workspace in-water robotic scanners for carefully controlled structural acoustics experiments. At the core of this capability is near-field acoustic holography (NAH), where dense near-field pressure measurements and Helmholtz based processing can be used to obtain a detailed mapping of surface pressure and normal velocity while eliminating diffraction and other edge effects. For the topics included in this paper, NAH measurements

were used only to investigate driver uniformity, and the results agreed with the laser Doppler vibrometer results shown previously.

The performance of the actuator is shown by the solid line of Fig. 2. The effective volume piezoelectric constant is found to be 400 pC/N. This high value is only 33% lower than the d_{33} constant of the material, indicating that the composite actuator material is performing quite well in reducing the influence of the negative d_{31}. The agreement between the measured and calculated response is quite good, particularly at low frequencies. At the higher frequencies shown there is some incidental reduction in output, most of which has been identified as being due to the addition of the other tile layers and components.

The measured free-field sensitivity of the pressure sensor (as assembled in the final tile) is shown in Fig. 5 for two evaluated configurations: with and without the lateral decoupler. Without the decoupler (dotted line) the sensitivity approaches the previously predicted value of -190 dB at low and high frequencies. However there is a substantial reduction in sensitivity in the frequency region where the lateral size of the sensor is between one-quarter and one-half the wavelength of sound in water. This is presumably due to the familiar "lateral mode" frequently observed in freely supported large area sensors, and hence is expected to be reduced when the tile is bonded to a backing structure. For the final configuration which includes the lateral decoupler (solid line), the measured sensitivity is nominally three decibels lower than predicted. This sensitivity reduction is more then
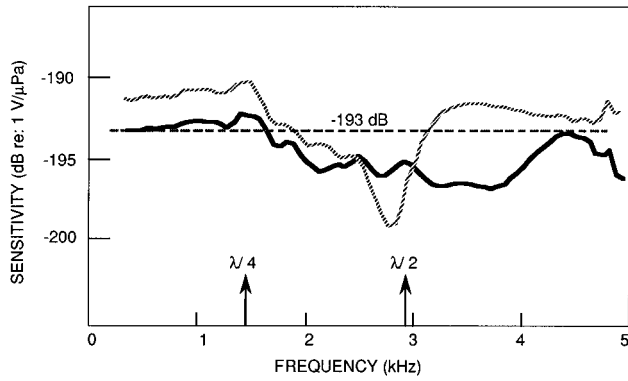
FIG. 5. Free-field sensitivity of the pressure sensor with (solid line) and without (dashed line) the lateral decoupler.

FIG. 7. Magnitude of the free-field sensor—actuator coupling. Dashed lines are predicted; pressure and velocity axis are displaced vertically 10 dB for clarity.

balanced in this application by the improved smoothness of the transfer function.

The measured and predicted free-field sensitivity of the velocimeter is shown in Fig. 6. The units used are equivalent free-field pressure, which is obtained by multiplying the velocity sensitivity by the acoustic impedance of water. This is convenient since it allows a direct comparison with the pressure sensor. (The measured sensitivity of each accelerometer in the velocity sensor is shown elsewhere.[5])

Also shown in Fig. 6 is the phase of the velocity sensor relative to that of the pressure sensor. The data for this tile construction show relatively little deviation from 180°, as expected (recalling that our velocity sign convention is for radiation).

C. Sensor—actuator coupling

The experimental measurement of pressure coupling C_p is straightforward. A known voltage is applied to the actuator, and the voltage obtained from the pressure sensor layer is measured. Coupling is then the ratio of the (complex) output to input voltage. The measured result is shown in Fig. 7. Also shown is the predicted value of the pressure coupling transfer function, where the calculation uses the previously

FIG. 6. Velocity sensor free-field sensitivity (equivalent free-field pressure units) and phase (re: pressure sensor).

calculated values for D and S_P, where the value of -193 dB re: 1 V/μPa is used for the later. The data are seen to be in agreement with the predicted result.

As a reference, the dashed line labeled ''infinite'' corresponds to the case where $\alpha_p = 1$. While this line is clearly not applicable to the single tile result, this curve is included here to illustrate the importance of properly including α_p. It is also worth noting that a large tile array with all tiles operating in unison would have a relative radiation impedance which would tend to approach unity, and the coupling would then approach this curve.

The velocity coupling is also shown in Fig. 7. Velocity coupling is measured by applying a voltage to the actuator, and measuring the average voltage obtained from the four accelerometers (in this case, prior to integration by the onboard electronic module). The measured magnitude of the velocity coupling is seen to be in agreement with that predicted (dashed line).

D. Combined sensor suitability

During the development of this tile it became apparent that our most critical and demanding test of sensor suitability involved combining of pressure and velocity sensor outputs to monitor tile reflectivity. Using a one-dimensional representation, for free-field insonification the reflection coefficient of the surface R is related to the impedance of the surrounding fluid Z_w and the acoustic input impedance of the surface Z_I which is simply the ratio of the surface pressure to the surface velocity. Then the surface reflection coefficient is given by

$$R = \frac{Z_w - Z_I}{Z_w + Z_I} = \frac{|p + \rho c v_n|}{|p - \rho c v_n|}, \tag{15}$$

where again the convention used in this case is that in which an outgoing velocity (radiation) is assigned a positive velocity value. As in the derivation of Eq. (1), the numerator represents the return or scattered field, while the denominator represents the incident field.

Corsaro *et al.*: Underwater surface impedance control   1579
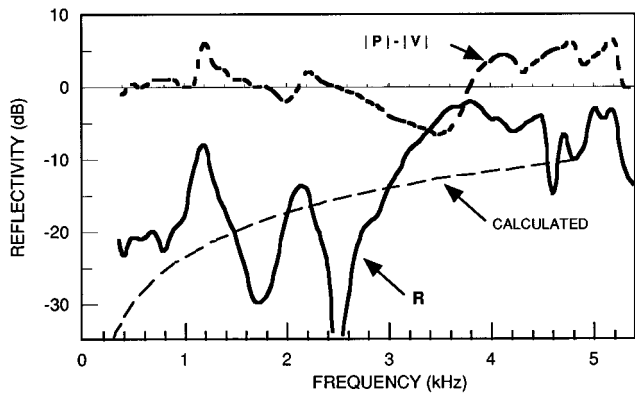
FIG. 8. Combined raw sensor outputs of preliminary tile design (without lateral decoupler).

FIG. 9. Combined raw sensor outputs of final tile as critical test of device suitability.

Hence an experimental measure (estimate) of this reflection coefficient can be obtained under far-field insonification conditions by forming the sum and difference of the pressure and velocity sensor outputs (after first amplifying the velocimeter output by a gain of ρc). This reflection coefficient can be compared to that calculated using appropriate layer properties.[16] For this calculation we represent the tile as a four layer structure and use a previously described numerical proceedure[4] based on the one-dimensional approach of Brekhovskikh.[17]

A comparison of the measured and predicted values provides an important test of the inherent ability of the tile to separate the incident and scattered fields. We consider this an essential test, even though control applications would certainly include the more nearly accurately measured frequency-dependent complex sensor calibration functions.

The usefulness of this test is illustrated in data of Fig. 8. The data shown were collected on a preliminary tile design prior to the inclusion of a the lateral decoupler. Pressure and velocity sensitivity and coupling data appeared smooth and well behaved, and in reasonable agreement with predictions. The difference between the pressure and velocity sensitivities was found to be typically less than 6 dB, as shown by the thick broken curve in Fig. 8. Smoothly varying sensitivity differences of this magnitude can typically be accommodated in a control system simply by including an accurate representation of transducer calibration in the processor function. The only early indication that the design might prove unsuitable in control applications occurred in examining the relative phase difference between the pressure and velocity sensors, but this in itself was inconclusive. However when the sensor outputs are combined to form the ratio given in Eq. (15), the resulting curve (solid line) is seen to differ markedly from the predicted reflection coefficient (dashed curve). Control systems analysis found that a tile design exhibiting this behavior would be difficult to use in a low-order robust control system.

The corresponding data for the final tile design are shown by the solid curve in Fig. 9. The agreement between expected and measured behavior is now quite good, with the only marked deviations occurring in the direction of reflectivity values which are lower than expected. This is encouraging, considering that the calculated curve does not include
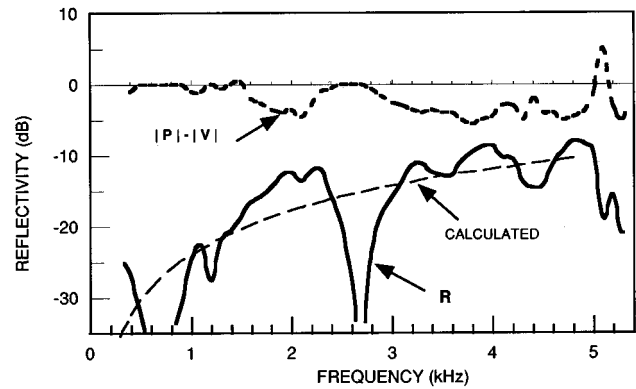
effects due to the finite size of the tile, and that the experimental curve is derived from raw transducer outputs without applying measured calibration functions or formal feedback control filter design. The only sensor information included in this test is a nominal pressure sensor sensitivity of -193 dB re: 1 V/μPa and an accelerometer sensitivity of 0.1 V/g. Control systems analysis subsequently confirmed the suitability of the transfer functions for this tile design for use with a low-order controller.[1]

V. CONCLUSIONS

The tile was specifically designed to be used in basic studies exploring the extent to which the acoustic characteristics of the tile surface can be interactively selected by the user via use of an active control system. To perform this function, the tile must be composed of transducer layers with well defined characteristics which contribute to a resulting integrated tile with well behaved and predictable performance characteristics.

As shown by the experimental data, the result of this effort is a tile with exceptionally favorable characteristics. The actuator has high authority and very good spatial and temporal uniformity. The sensors have high uniform response with wide dynamic range and low noise figures. The models used in the design of this tile appear to have been generally successful in describing the measured acoustic transfer functions. The mechanisms responsible for most features of the data have been identified. The combined dynamic operating range of the integrated tile is quite wide—typically 70–90 dB (depending on frequency). Finally, the coupling between actuator and sensors is well behaved, relatively uniform, and typically in the 60–80 dB range. Since this coupling figure is related to the neutralization gain requirement, it suggests that the control system required to form a smart actuator system need only have relatively low gain requirements, which implies stability and robustness.

ACKNOWLEDGMENTS

[1] R. D. Corsaro and B. H. Houston, ''Sensor-actuator interactions in the NRL-ABC platform,'' J. Acoust. Soc. Am. **98**, 2903 (1995).

[2] A. Parvalescu, ''Active electroacoustic networks and the synthetic absorber,'' J. Acoust. Soc. Am. **33**, 1668 (1961).

[3] D. M. Photiadis, J. A. Bucaro, and R. D. Corsaro, ''Double layer actuator,'' J. Acoust. Soc. Am. **96**, 1613–1619 (1994).

[4] R. D. Corsaro and R. M. Young, ''Influence of backing compliance on transducer performance,'' J. Acoust. Soc. Am. **97**, 2849–2854 (1995).

[5] R. D. Corsaro and B. H. Houston, ''Acoustic velocity sensor for the NRL ABC research platform,'' in *Acoustic Particle Velocity Sensors: Design, Performance, and Applications*, edited by M. J. Berliner and J. F. Lindburg, AIP Conference Proceedings 368 (AIP, Woodbury, NY, 1996).

[6] Product literature from Material Systems, Inc. (Littleton, MA).

[7] L. E. Kinsler, A. R. Frey, A. Coppens, and J. Sanders, *Fundamentals of Acoustics*, 3rd ed. (Wiley, New York, 1982), pp. 176–177.

[8] A. C. Tims, ''Hydrophone preamplifier optimization prediction of hydrophone self-noise by a noise model,'' NRL Report 8180 (1978).

[9] Product literature from Ithaco, Inc. (Ithaca, NY), adjusted using laboratory measurements at NRL on the particular unit used.

[10] M. C. Junger and D. Feit, *Sound, Structures, and Their Interactions*, 2nd ed. (MIT Press, Cambridge, MA, 1982), p. 211.

[11] Product literature from Wilcoxon Research, Inc. (Gaithersburg, MD).

[12] T. Christian and G. Corcoran, NRL Memorandum Report.

[13] J. A. Archer-Hall and D. Gee, ''A single integral computer method for axisymmetric transducers with various boundary conditions,'' NDT Int. **13**, 95–101 (1980).

[14] N. Yen (Naval Research Lab.), private communication.

[15] M. Marcus (Naval Research Lab.), private communication.

[16] R. D. Corsaro, in *Sound and Vibration Damping with Polymers*, edited by R. D. Corsaro and L. H. Sperling (American Chemical Society, Washington, D.C., 1990).

[17] L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1960), pp. 45–60.

Harmonic distortion in silicon condenser microphones

Michael Pedersen,[a] Wouter Olthuis, and Piet Bergveld

MESA Research Institute, University of Twente, P.O. Box 217, NL-7500 AE Enschede, The Netherlands

The nonlinear effects causing distortion in condenser microphones are investigated and modeled with special attention paid to microphones made with silicon micromachining technology. A model of the distortion, based on the quasistatic properties, is described and solved numerically using finite differences. The results of the model show that the distortion is mainly caused by both the inverse relation between the microphone capacitance and the deflection of the diaphragm, as well as the nonlinear electrostatic forces generated by the applied dc bias voltage. The results of the model are compared with measurement results on a silicon/polyimide microphone, and good agreement is found. © *1997 Acoustical Society of America.* [S0001-4966(97)01809-2]

PACS numbers: 43.38.Kb, 43.58.Ry [SLE]

INTRODUCTION

The further miniaturization of capacitive microphones by the introduction of silicon microfabrication technologies has led to devices[1–6] in which the dc bias voltage required to generate the electrical field in the microphone has been reduced significantly, while maintaining a high performance of the microphone. This has been achieved by reducing the air gap between the diaphragm and the backplate, which can be reproducibly made with the silicon technology in the range of 1–5 μm (Fig. 1). However, in the quest for high sensitivities and low noise levels, the overall size, thickness, and therefore the mechanical sensitivity, of the diaphragm are similar to conventional miniature condenser microphones. The result of this is that the nonlinear effects in the transducer become evident at a lower sound-pressure level (SPL), as the deflection of the diaphragm compared to the air gap is much larger than for conventional devices with air gaps in the range of 15 to 30 μm. The nonlinearity in the capacitive microphone has two major contributions. First, the nonlinear and nonuniform electrostatic attraction forces between the diaphragm and the backplate, generated by the dc bias voltage, have influence on the distortion. For bias voltages close to the collapse of the structure, this is the dominating source of distortion. Second, the dynamic deflection of the diaphragm causes a dynamic change of the electrical field between the diaphragm and the backplate, leading to an asymmetric dynamic response.

The issue of distortion in silicon condenser microphones has until now been treated with little attention, and few measurements have been reported.[3] In this paper we present a general theory and simulations of the harmonic distortion in capacitive microphones, and the results of the analysis are compared with measurements on condenser microphones. These condenser microphones have been fabricated using a silicon surface micromachining process based on polymer processing.[7]

---

[a] Electronic mail: psn@eltn.utwente.nl

I. THEORY AND SIMULATION OF HARMONIC DISTORTION IN CAPACITIVE MICROPHONES

Harmonic distortion in transducers is in general caused by the nonlinearity always present in the device. This nonlinearity is normally due to a combination of the exact nature of the transduction principle and more intrinsic nonlinear effects of the material parameters. A common method to describe these phenomena is to analyze the frequency power spectrum of the output of the transducer with a harmonic input signal. Due to the nonlinear behavior of the transducer, harmonics of higher order will be generated on the output. By dividing the amplitude of these higher harmonics by the amplitude of the output signal generated with the fundamental angular frequency ω of the input signal, the total harmonic distortion of the transducer may be calculated:

$$\text{THD} = \sqrt{\frac{V_{2\omega}^2 + V_{3\omega}^2 + V_{4\omega}^2 + \ldots + V_{n\omega}^2}{V_\omega^2}}. \qquad (1)$$

In the particular case of condenser microphones, there are different approaches to model the distortion. One approach is to describe the electromechanical problem by means of the analogy between the different physical domains,[8] which can be used to derive the transducer equations including the nonlinear attraction forces from the dc bias voltage.[9] One of the limitations of this theory is that the diaphragm (and backplate) are supposed to move in a piston-like manner. Considering the microphone in Fig. 1, this is clearly not the case and consequently the nonuniformity of the electrostatic attraction forces cannot be modeled. A solution which includes a more realistic shape of the diaphragm deflection has been applied in the calculation of static and dynamic diaphragm deflections by Busch-Vishniac,[10] who assumed a series solution for the deflection of a circular diaphragm in an electrostatic earphone. From this solution, a leading order distortion of the response could be derived. An alternative to the analytical solutions is to solve the problem numerically. Numerical methods have previously been applied to the simulation of the static and dynamic responses of condenser microphones,[11–13] and will in the following be applied to calculate the harmonic distortion.
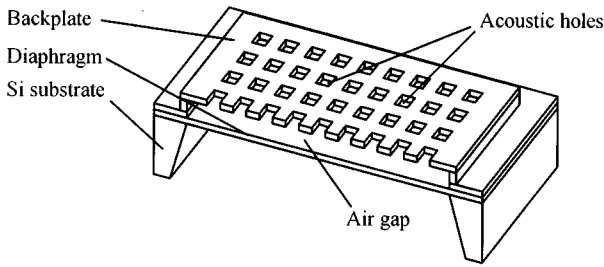
FIG. 1. Cross-sectional view of a typical condenser microphone made with silicon micromachining technology.

Assuming the frequency of operation is far below resonance of the diaphragm or roll-off of the microphone, the mass of inertia and the damping terms may be omitted, and the dynamic deflection of the square diaphragm under influence of both bending and built-in stress is defined by the following quasistatic equation:

$$\frac{E_d h_d^3}{12(1-\nu_d^2)}\left(\frac{\partial^4 w_d}{\partial x^4}+2\frac{\partial^4 w_d}{\partial x^2 \partial y^2}+\frac{\partial^4 w_d}{\partial y^4}\right)$$

$$=q_{\mathrm{sp}}+q_{\mathrm{el}}+\sigma_d h_d\left(\frac{\partial^2 w_d}{\partial x^2}+\frac{\partial^2 w_d}{\partial y^2}\right), \quad (2)$$

where E_d is the Young's modulus, ν_d is Poisson's ratio, w_d is the two dimensional deflection function, σ_d is the built-in stress, h_d is the thickness, q_{sp} is the quasistatic sound pressure, and q_{el} is electrostatic force per unit area generated by the applied dc bias voltage. In contrast to conventional condenser microphones, the backplate in silicon condenser microphones generally has a thickness comparable to the diaphragm. Therefore, the deflection of the backplate cannot be neglected, and can be described by an equation similar to (2):

$$\frac{h_b^3}{12}\left[C_1\left(\frac{\partial^4 w_b}{\partial x^4}+\frac{\partial^4 w_b}{\partial y^4}\right)+(2C_2+4C_3)\frac{\partial^4 w_b}{\partial x^2 \partial y^2}\right]$$

$$=q_{\mathrm{el}}+\sigma_b h_b C_4\left(\frac{\partial^2 w_b}{\partial x^2}+\frac{\partial^2 w_b}{\partial y^2}\right), \quad (3)$$

in which h_b, σ_b, and w_b are the thickness, built-in stress, and the deflection function of the backplate. The constants C_1–C_4 denote the reduced rigidity of the backplate due to the presence of the acoustic holes. These constants have been derived for a regular pattern of square holes with side length $2a$ and center-to-center distance b using the strain energy equalization method:[14]

$$C_1=\frac{E_b}{b^2}\left[\frac{b(b-2a)}{1-\nu_b^2}+\frac{a(b-2a)^2}{b}\right],$$

$$(4)$$

$$C_2=\frac{\nu_b E_b(b-2a)}{b(1-\nu_b^2)},$$

$$C_3=\frac{E_b}{4b^2(1+\nu_b)}\left[2b(b-2a)+\frac{12Na(b-2a)}{bh_b^3}\right],$$

$$(5)$$

$$C_4=\sqrt{1-\frac{4a^2}{b^2}},$$

where N is a constant depending on the exact dimensions of the holes in the backplate, related to the thickness.[14]

The nonlinear nonuniform attraction force per unit area between the diaphragm and the backplate, which is generated by the dc bias voltage, is expressed by

$$q_{\mathrm{el}}=K\frac{\varepsilon_d \varepsilon_b \varepsilon_0}{2[\varepsilon_d h_b+\varepsilon_b h_d+\varepsilon_d \varepsilon_b(h_a-w_d-w_b)]^2}V_b^2, \quad (6)$$

in which ε_0 is the vacuum permittivity, ε_d and ε_b are relative permittivities of the diaphragm and the backplate, h_a is the initial air gap between the diaphragm and the backplate, V_b is the dc bias voltage, and K is a factor accounting for the reduction of the electrostatic forces due to the acoustic holes in the backplate. The sign convention of w_d and w_b is such that deflections are positive toward the center of the air gap.

With Eqs. (2), (3), and (6), a quasistatic model of the microphone structure has been derived. This model is particularly useful in determining the stability of the condenser microphone, since it can be used to analyze the nonlinear relation between the sensitivity and the applied bias voltage. This relation must be known for a specific structure to fully optimize the performance of the microphone. As it can be seen, the deflection functions w_d and w_b both appear in the expression for the electrostatic load q_{el}, thereby posing a complex mathematical problem in the solution of (2) and (3). Furthermore, the equation system can become unstable due to the inverse relation between q_{el} and w_d and w_b. In this situation, no solution to the system exists, which may be interpreted as a collapse of the structure. The values of V_b and q_{sp} where this collapse occurs are known as the critical bias voltage and the maximum allowable sound pressure, and are the two most important stability parameters of the condenser microphone.

Unfortunately, due to the complexity of the equation system, no analytical solutions to w_d and w_b are known, and consequently a numerical simulation method must be employed. For this purpose, an iterative numerical method based on finite differences has been developed and implemented to solve the two nonlinear coupled equations.[12] With this method, the continuous equation system (2), (3), and (6) is replaced by the discrete system:

$$\overline{\overline{A}}_d\cdot\overline{w}_d=\overline{q}_{\mathrm{sp}}+\overline{q}_{\mathrm{el}}$$

$$\overline{\overline{A}}_b\cdot\overline{w}_b=\overline{q}_{\mathrm{el}} \quad (7)$$

$$\overline{q}_{\mathrm{el}}=f(\overline{w}_d,\overline{w}_b),$$

where $\overline{\overline{A}}_d$ and $\overline{\overline{A}}_b$ are coefficient matrices of (2) and (3), and f is the discrete function of (6). Moreover, a quasistatic sensitivity of the microphone can be calculated by solving the equation system for two different sound pressures q_{sp1} and q_{sp2}. The change of capacitance ΔC due to the change of sound pressure can be found by integration of the deflection
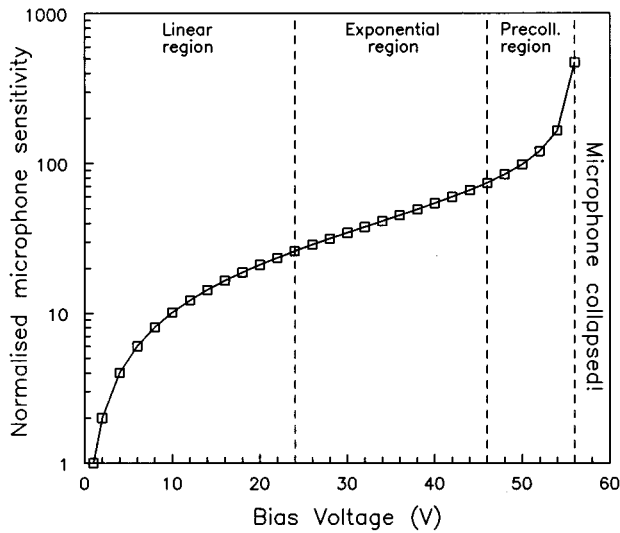
FIG. 2. Typical relation between sensitivity and dc bias voltage of a condenser microphone.

functions over the total sensitive microphone area according to

$$\Delta C = \int \int \frac{\varepsilon_0}{\dfrac{h_b}{\varepsilon_b} + \dfrac{h_d}{\varepsilon_d} + h_a - w_{d1} - w_{b1}} \, dx \, dy$$

$$- \int \int \frac{\varepsilon_0}{\dfrac{h_b}{\varepsilon_b} + \dfrac{h_d}{\varepsilon_d} + h_a - w_{d2} - w_{b2}} \, dx \, dy. \qquad (8)$$

The open-circuit quasistatic sensitivity may be calculated from

$$S = \left| \frac{\Delta C}{q_{sp1} - q_{sp2}} \right| \cdot \frac{V_b}{C_0}, \qquad (9)$$

in which C_0 is the capacitance of the microphone without sound pressure applied.

   With these equations, the dependence of the sensitivity on the bias voltage and SPL can be thoroughly analyzed. The result of this kind of analysis is a curve with the form shown in Fig. 2. This curve is typical for condenser microphones, and can be naturally split in four different regions. In the first linear region for small bias voltages, the electrostatic forces are insignificant compared with the acoustic load, and the sensitivity increases linearly with the electric field in the microphone. For intermediate voltages, the electrostatic forces have a notable influence, causing an exponential dependence of the sensitivity, and in the region close to collapse of the structure, the influence becomes significant enough to cause a dependence stronger than exponential. This dependence on the dc bias voltage has been shown to be in good agreement with measurements far below the roll-off frequency of the microphone.[14]

   The quasistatic model can furthermore be used to simulate the distortion of the microphone, providing that the frequency of operation is far below the roll-off frequency or the resonance frequency of the diaphragm. In this situation, the

response of the microphone may be calculated by solving the equation system for each value of the acoustic sound pressure. The applied harmonic acoustic sound pressure is assumed to have the form

$$q_{sp}(t) = q_A \sin(\omega t). \qquad (10)$$

The microphone capacitance may be calculated by integration similar to (8):

$$C_m(t) = \int \int \frac{\varepsilon_0}{\dfrac{h_b}{\varepsilon_b} + \dfrac{h_d}{\varepsilon_d} + h_a - w_d(t) - w_b(t)} \, dx \, dy, \qquad (11)$$

where the time-dependent deflection functions w_d and w_b of the diaphragm and the backplate must be found by solving the equation system (2), (3), and (6). Assuming that the electrical field remains constant in the air gap of the microphone, the open-circuit response of the microphone can be calculated by use of the basic capacitor equation:

$$Q = C_m \cdot V \quad \frac{\partial Q}{\partial t} = \frac{\partial C_m}{\partial t} V + C_m \frac{\partial V}{\partial t} = 0. \qquad (12)$$

This simple differential equation for V has the following solution:

$$V_o(t) = K_l \exp\left(-\int \frac{\partial C_m}{\partial t} \frac{1}{C_m} dt \right), \qquad (13)$$

in which K_l is a constant of integration used to fulfill the boundary condition $V_o(0) = V_b$. With the previous definitions of C_m, the output signal can be calculated according to (13).

   The assumption of a constant electrical field is valid only when the deflections of the diaphragm are small compared to the air gap. For larger diaphragm deflections, the coupling of the electrical field to the dynamic change of the air gap must be included. The output signal of the microphone is defined as electrical field $(E) \times$ diaphragm displacement (d), which with the definitions above leads to

$$V_o(t) = E(t) \cdot d(t) = \frac{V_b}{h_{eff} - w_{d,c}(t) - w_{b,c}(t)}$$
$$\times [w_{d,c}(t) + w_{b,c}(t)], \qquad (14)$$

where $w_{d,c}$ and $w_{b,c}$ represent the center deflections of the diaphragm and the backplate, which can be calculated with the equation system (2), (3), and (6), and h_{eff} is the effective air gap given by

$$h_{eff} = h_a - w_{d,c0} - w_{b,c0}, \qquad (15)$$

in which $w_{d,c0}$ and $w_{b,c0}$ are the center deflections induced only by the electrostatic forces (zero sound pressure). Using the deflections in the center of the diaphragm and backplate gives a worst-case estimation of the distortion, since the dynamic change of the air gap is largest in the center. In Fig. 3, the calculated output signals from (13) and (14) are shown for a condenser microphone with the dimensions and material parameters given in Table I. The resonance frequencies of the diaphragm and the backplate have been calculated using a lumped element model[7] and were estimated to be 91
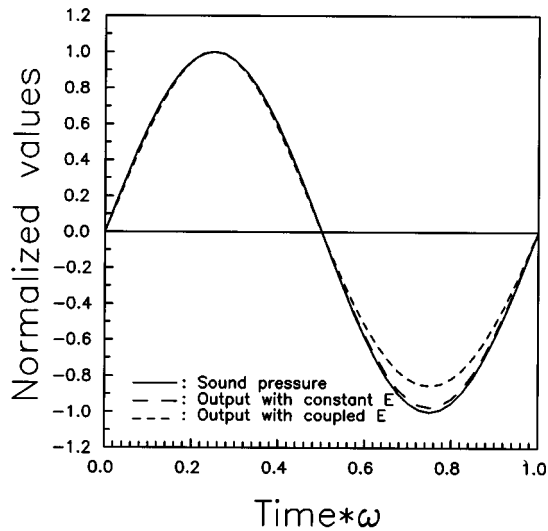
FIG. 3. Simulated microphone output voltages (dashed lines) for harmonic sound pressure (solid line).

FIG. 4. Simulated THD vs SPL for constant (dashed line) and coupled (solid line) electrical field (dc bias voltage: 10 V).

and 100 kHz, respectively. A harmonic sound pressure with amplitude $q_A = 125$ dB (re: 20 μPa) was applied (solid line), and the output voltages (dashed lines) were simulated with a dc bias voltage $V_b = 10$ V. As can be seen from Fig. 3, both calculated output voltages are asymmetric, and the output voltage from (14) is more distorted than that of (13). This is further supported by the results shown in Fig. 4, in which the distortion of the first five harmonics was calculated as a function of the applied SPL for a dc bias voltage of 10 V. The indication of these results is that the one directional electrostatic forces, represented both in (13) and (14), only provide a small contribution to the overall distortion, and that the dynamic change of the electrical field, included in (14), dominates the distortion. It can be seen from Fig. 4 that the distortion follows a power-law for increasing SPLs. This is in good agreement with other results,[15] and means that the distortion depends linearly on the sound pressure. With increasing SPLs the distortion increases rapidly, as the sound pressure approaches the maximum allowable sound pressure, where the microphone structure collapses. Moreover, an increase of the distortion is detected for higher dc bias voltages, due to the increasing electrostatic forces (Fig. 5). The electrostatic forces contribute to the distortion in two manners. First, a static deflection of the diaphragm and the back-

plate is induced, thereby reducing the effective air gap in the microphone. This leads to a larger dynamic change of the electrical field according to (14). Second, the electrostatic forces increase the movements of the diaphragm toward the backplate, and attenuate the movements away from the backplate, thereby adding further to the asymmetry of the output signal.

II. EXPERIMENTAL RESULTS

Condenser microphones, with the dimensions given in Table I, and with polyimide diaphragms and backplates, have been fabricated using a silicon micromachining process.[7] In Fig. 6 a secondary emission monitor (SEM) photograph is shown, depicting the front of the microphone with the perforated polyimide backplate, and two bond pads for electrical connection.

TABLE I. Dimensions and parameters of the simulated and implemented microphone structure.

	Polyimide diaphragm	Polyimide backplate
Young's modulus	3 GPa	3 GPa
Poisson's ratio	0.3	0.3
Built-in stress	70 MPa	50 MPa
Relative permittivity	3.4	∞[a]
Thickness	1.0 μm	17 μm
Air gap	1.5 μm	
Side length	1.6 mm	
Acoustic hole fraction	36 % ($2a$: 75 μm, b: 125 μm)	

[a]The electrode is placed under the backplate.

FIG. 5. Simulated THD vs SPL and dc bias voltage of the condenser microphone (Table I).
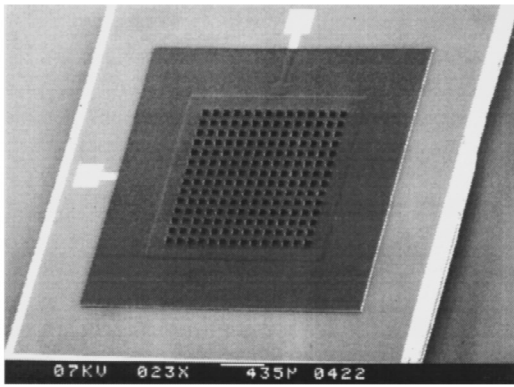
FIG. 6. SEM photograph of the front of the polyimide microphone.

FIG. 8. Measured (symbols) and simulated (lines), THD vs SPL.

The microphones were characterized using a pressure type acoustical test station, which has been described elsewhere.[16] The method used for acoustical characterization is illustrated in Fig. 7. The loudspeaker is driven by a Hewlett–Packard 35670A dynamic signal analyzer, which also measures the responses of the reference microphone and the microphone under test (MUT). To measure the distortion, the loudspeaker is excited with a 1-kHz sine, the amplitude of which is adjusted to generate the specified SPL. The reference microphone is used in this case to measure the SPL and the distortion of the loudspeaker. Due to limitations of the loudspeaker, sound pressure of up to 125 dB SPL may be produced without significant distortion ($<0.06\%$). At such levels, the distortion of the reference microphone itself is negligible. The total harmonic distortion is determined with a built-in function in the analyzer, which calculates the weight of the higher harmonics in the power spectrum to the amplitude at the fundamental frequency. In Fig. 8, the measured distortion is shown as a function of the SPL for two different values of the dc bias voltage. Also shown are the simulated values according to (14), and it can be seen that excellent agreement exists between the theory and the measurements. The deviations at low SPLs (<100 dB re: 20 μPa) are due to the limited signal-to-noise ratio of the measurement setup. The relation between distortion and dc bias voltage has also been measured (Fig. 9) for an SPL of 120 dB (re: 20 μPa). Again, very good agreement between measurements and simulations is found; however, for bias voltages close to the collapse of the microphone, the worst-case model overestimates the distortion. The measured distortion for a dc bias voltage of 10 V was 2.17%, compared to a simulated value of 2.20%.

III. CONCLUSIONS

In this paper a theoretical model for the simulation of harmonic distortion in condenser microphones has been proposed. The model was derived by extending a quasistatic model, which has previously been described. The simulation of the distortion, which is a dynamic problem, is valid with this quasistatic method for frequencies far below the roll-off frequency of the microphone. The distortion in the microphone was found to be linearly dependent on the sound pressure for conditions below collapse. Furthermore, the relation between the distortion and dc bias voltage was found to be a strongly nonlinear function, which increases sharply as the microphone approaches the collapsed situation. The theoret

FIG. 9. Measured (symbols) and simulated (lines), THD vs dc bias voltage (SPL: 120 dB re: 20 μPa).

FIG. 7. Measurement of the acoustical microphone response.

Pedersen *et al.*: Harmonic distortion in microphones

ical model was compared to distortion measurements performed on micromachined condenser microphones. The results of these measurements were in correlation with the predictions of the model, and therefore it may be concluded that the model is sufficiently realistic to provide information about the complex physical situation in the microphone structure. Considering the relation between the distortion and the dc bias voltage, it can be concluded that harmonic distortion may be one of the limiting parameters regarding the performance, since a dc bias voltage (sensitivity) smaller than the highest possible (optimum) may have to be chosen to keep the distortion within reasonable magnitude. It is believed that the model of the distortion can be useful in the design phase of capacitive microphones, where the optimum balance between performance and stability must be found.

A copy of the computer code used to attain the results described in this paper may be acquired with operating instructions from the authors.

ACKNOWLEDGMENTS

[1] D. Hohm and R. Gerhard-Multhaupt, ''Silicon-dioxide electret transducer,'' J. Acoust. Soc. Am. **75**, 1297–1298 (1984).

[2] D. Hohm and G. Hess, ''A subminiature condenser microphone with silicon nitride membrane and silicon back plate,'' J. Acoust. Soc. Am. **85**, 476–480 (1989).

[3] J. Bergqvist, F. Rudolf, J. Maisano, F. Parodi, and M. Rossi, ''A silicon condenser microphone with a highly perforated backplate,'' in *Proceedings of the 6th International Conference on Solid-State Sensors and Actuators*, San Francisco, June 1991, pp. 266–269.

[4] W. Kühnel and G. Hess, ''A silicon condenser microphone with structured back plate and silicon nitride membrane,'' Sens. Actuators A **30**, 251–258 (1992).

[5] J. Bergqvist and F. Rudolf, ''A silicon condenser microphone using bond- and etch-back technology,'' Sens. Actuators A **45**, 115–124 (1994).

[6] P. R. Scheeper, W. Olthuis, and P. Bergveld, ''Improvement of the performance of microphones with a silicon nitride diaphragm and backplate,'' Sens. Actuators A **40**, 179–186 (1994).

[7] M. Pedersen, W. Olthius, and P. Bergveld, ''A silicon condenser microphone with polyimide diaphragm and backplate,'' Sens. Actuators A (to be published).

[8] H. F. Olson, *Acoustical Engineering* (Van Nostrand, Princeton, 1957).

[9] C. H. Sherman and J. L. Butler, ''Analysis of harmonic distortion in electroacoustic transducers,'' J. Acoust. Soc. Am. **98**, 1596–1611 (1995).

[10] I. J. Busch-Vishniac, ''Response of an edge-supported circular membrane electret earphone. Part I—Theory,'' J. Acoust. Soc. Am. **75**, 977–989 (1984).

[11] J. E. Warren, A. M. Bezezinski, and J. F. Hamilton, ''Capacitance-microphone static membrane deflections,'' J. Acoust. Soc. Am. **52**, 711–719 (1972).

[12] M. Pedersen, W. Olthuis, and P. Bergveld, ''On the simulation of subminiature condenser microphones using finite differences,'' in *Proceedings of the 1st International Conference on the Simulation and Design of Microsystems and Microstructures, Southampton, UK, September* 1995, pp. 185–192.

[13] J. E. Warren, A. M. Bezezinski, and J. F. Hamilton, ''Capacitance microphone dynamic membrane deflections,'' J. Acoust. Soc. Am. **54**, 1201–1213 (1973).

[14] M. Pedersen, W. Olthuis, and P. Bergveld, ''On the mechanical behaviour of thin perforated plates and their application in silicon condenser microphones,'' Sens. Actuators A **54**, 499–504 (1996).

[15] Data Handbook, Condenser microphones and microphone preamplifiers for acoustic measurements (revision September 1982, Brüel & Kjær).

[16] M. Pedersen, R. Schellin, W. Olthuis, and P. Bergveld, ''Electro-acoustical measurements of silicon microphones on wafer scale,'' J. Acoust. Soc. Am. **101**, 2122–2128 (1997).

Active structural acoustic control using strain sensing

P. Masson, A. Berry, and J. Nicolas

G.A.U.S., Mechanical Engineering Department, Université de Sherbrooke, Sherbrooke, Québec J1K 2R1, Canada

A new strategy based on the direct use of strain sensing is proposed for the minimization of the acoustic radiation from a vibrating plate. This strategy uses active structural acoustic control because it eliminates the need for far-field acoustic sensors. The cost function is defined as the radiated acoustic power and the discrete strain information is given directly, for example, by fiber optic sensors embedded in the structure. Two approaches are presented to express the cost function in terms of strain information. The first one is based on plate's displacement reconstruction from the strain field while the second is based on integration by parts of the wave-number transform of the displacement. Proper concentration of the control effort is obtained with the cost function defined in the wave-number domain. Both approaches are first validated with respect to an analytical solution for the calculation of the radiated acoustic power from structural strain information. It is shown that good agreement is obtained with at least two strain sensors per structural wavelength. Optimal control of acoustic radiation from a simply supported plate is then performed using different configurations of sensors and actuators. The control is performed using simulated piezoelectric actuators and results show that both approaches are well suited for active control applications as they offer good performance over a wide frequency range while requiring only few sensors. © *1997 Acoustical Society of America.* [S0001-4966(97)03509-1]

INTRODUCTION

Active control of acoustic radiation from vibrating structures has been investigated by many researchers.[1] Among the proposed strategies, the approaches which control the radiation by actively acting on the structure have been referred to as active structural acoustic control (ASAC).[2]

Active control of radiation is considered in situations where low-frequency acoustic fields have to be controlled.[3] ASAC has been demonstrated to have a lower requirement on the number of secondary sources than entirely acoustic strategies controlling directly the acoustic field with loudspeakers.[1]

ASAC strategies make use of three components: error sensors, which define the type of information to be minimized; a minimization algorithm, which is closely linked to a properly defined cost function; and a set of actuators used as secondary sources to actively interact with the structure. Error sensors can be placed in the acoustic field to be controlled (microphones) or directly on the radiating structure. The actual trend towards smart materials calls for the integration of the error sensors into the structure. Following the same idea, most of the actuators [shape-memory alloys, electrorheological fluids, polyvinylidene fluoride films (PVDF), piezoelectric ceramics (PZT), etc.] used are now designed to be embedded in the structure.[4–7]

This paper is thus concerned with the definition of the cost function relating a sensing information to a quantity to be minimized, in order to achieve adequate control. This strategy is then assessed in the context of optimal control of acoustic radiation from a simply supported plate so that practical implementation can be evaluated.

Several techniques are available to estimate the acoustic radiation from vibrating structures. Among them, intensity or pressure measurements can be used to evaluate the radiated acoustic power.[8] Acoustic holography has also been applied to the prediction of far-field radiation.[9] However, the most promising sensing strategies used in ASAC involve monitoring of acceleration, velocity, or displacement of the structure, typically by means of accelerometers,[10–13] or polyvinylidene fluoride films (PVDF)[14–20] for strain monitoring. These approaches eliminate the need for far-field acoustic sensors, following the tendency towards complete integration in the structure. The radiation of the structure is then estimated either directly[21] or from a model of the radiating structure.[22] The direct approach has the important advantage of not being dependent on the accuracy of any model.

In this paper, a new sensing strategy is presented in which the estimation of the acoustic radiation involves monitoring the strain field at discrete points on the structure. Such information will be ultimately given by fiber optic sensors, for example, which can be integrated in the structure. Moreover, as the sensing is not based on any predetermined model, it is straightforward to make the control fully adaptive. It was previously shown that fiber optic sensors can be used in active control as distributed strain sensors.[23,24]

Two approaches are presented within the proposed strategy, using the second spatial derivative of the displacement (or strain) information. The ASAC is performed in the wave-number domain to take advantage of physical mechanisms[25,26,21] and the cost fuction is defined as the radiated acoustic power so that global control is achieved. It appears that the wave-number domain is a natural one for understanding structural radiation. The strategy is oriented towards application to the active control of the acoustic radiation
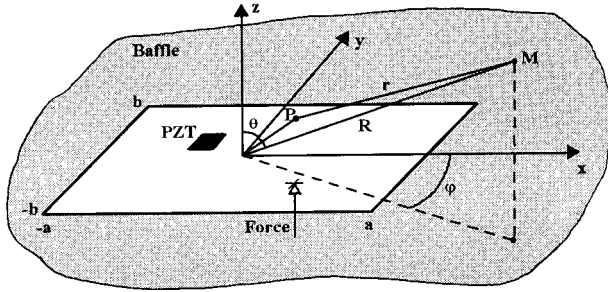
FIG. 1. Coordinate system for the rectangular plate.

from a baffled rectangular plate. While strain sensing using PVDF suffers from ambiguous measurements on two-dimensional radiating structures (all the strain components are included in the signal), the use of discrete fiber optics strain sensing permits the discrimination of the components of the strain, thus allowing proper wave-number domain control.

The first part of the paper presents theoretical developments leading to the definition of a cost function using two approaches: The first one is based on a finite differences displacement reconstruction scheme and the second one uses a double integration by parts of the wave-number transform of the displacement. In the second part, the performances of both approaches are compared in the context of optimal control. A simply supported plate is excited by a point force disturbance and the control is performed by simulating the effect of piezoelectric patches (PZT) as actuators. The disturbance and actuators equations pertaining to the case of a simply supported plate are presented and optimal control results are presented using two actuating (PZT) configurations.

I. RADIATED ACOUSTIC POWER FROM STRAIN INFORMATION

The cost function to be used in active control may be defined in different ways. The acoustic pressure radiated in prescribed directions has been used by some researchers.[11–13] However, in order to ensure proper global acoustic control, the radiated acoustic power is an appropriate quantity to be minimized.[8]

The acoustic radiation from a baffled rectangular thin plate in a fluid is considered in this work. The geometrical situation is presented in Fig. 1. A point force disturbance is used to excite the structure at a given frequency ω_0.

A. Acoustic radiation

The acoustic pressure $p(M)$ radiated at a point M from a baffled rectangular plate can be expressed by the Rayleigh's integral:

$$p(M) = \frac{\rho_0}{2\pi} \int_A \ddot{W}(P) \frac{e^{-ikr}}{r} \, dA, \qquad (1)$$

where ρ_0 is the mass density of the fluid, $\ddot{W}(P)$ is the second time derivative of the plate's displacement at a point P, k is the acoustic wave number in the fluid, r is the distance between the point corresponding to the displacement consid-

ered and the reception point in the fluid, and A is the surface of the plate. Assuming harmonic displacement such that $W(P,t) = w(x,y)e^{i\omega t}$ the preceding equation can be used to obtain a far-field approximation to the radiated acoustic pressure:[27]

$$p(R,\lambda,\mu) = -\frac{\omega^2 \rho_0}{2\pi} \frac{e^{-ikR}}{R} \tilde{w}(\lambda,\mu), \qquad (2)$$

where

$$\lambda = k \sin \theta \cos \varphi, \qquad (3)$$

$$\mu = k \sin \theta \sin \varphi \qquad (4)$$

and where R, θ, and φ are the spherical coordinates of the point M, ω is the angular frequency, $\tilde{w}(\lambda,\mu)$ is the double space-wave-number transform of the plate's displacement, analogous to the well-known time-frequency Fourier transform, and defined as:[26]

$$\tilde{w}(k_x,k_y) = \int_{-a}^{a} \int_{-b}^{b} w(x,y)e^{i(k_x x + k_y y)} \, dy \, dx \qquad (5)$$

and k_x and k_y are, respectively, the x and y components of the structural wave number.

The acoustic power Π radiated from the plate can be expressed using the integral over a hemisphere of the mean square of the pressure in the far field:

$$\Pi = \frac{1}{2\rho_0 c} \int_S p(R,\theta,\varphi)p^*(R,\theta,\varphi)dS, \qquad (6)$$

where S is the surface of the hemisphere, c is the speed of sound in the fluid, and the asterisk denotes the complex conjugate.

B. Radiated acoustic power in the wave-number domain

In order to exploit the physical mechanisms of structural radiation, the cost function will be defined in the wave-number domain.[25,28,26] To allow the definition of such a cost function, Eq. (6) is rewritten in the wave-number domain:

$$\Pi = \frac{R^2}{2\rho_0 \omega} \int_{-k}^{k} \int_{-\sqrt{k^2-\lambda^2}}^{\sqrt{k^2-\lambda^2}} \frac{p(R,\lambda,\mu)p^*(R,\lambda,\mu)}{\sqrt{k^2-\lambda^2-\mu^2}}$$

$$\times d\mu \, d\lambda. \qquad (7)$$

In Eq. (7), the integration is performed over the supersonic region of the structural wave-number spectrum where the acoustic wave number k is larger than the structural wave number given by $\sqrt{\lambda^2+\mu^2}$, which shows that only supersonic structural waves radiate sound in the far field. The minimization of the supersonic region of the displacement wave-number spectrum will then minimize the radiated acoustic power. Equation (7) relates the radiated acoustic power to the far-field pressure and will be used as the cost function in the active control scheme.

1589   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Masson *et al.*: Strain sensing in active control   1589
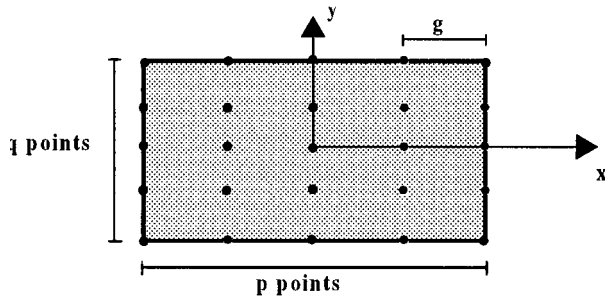
FIG. 2. Strain measurement points.

C. Acoustic pressure from discrete strain information

The acoustic pressures appearing in Eq. (7) can be written in terms of the discrete strain information. Under the Kirchhoff thin plate theory, the strain ϵ_i at the surface of the plate, in a particular direction ξ_i, is related to the second space derivative, in that direction, of the displacement field $w(x,y)$:[29]

$$\epsilon_i = \frac{h}{2}\frac{\partial^2 w(x,y)}{\partial \xi_i^2}, \tag{8}$$

where h is the thickness of the plate.

It is assumed that the strain information is known only at an arbitrary number of points on an equally spaced $p \times q$ grid of sensors. Figure 2 presents schematically the sampling grid for a 5×5 array of strain sensors. The Nyquist criterion also applies to the space-wave-number transform and it is therefore expected that a spatial sampling rate of two measurement points per shortest structural wavelength present in the signal will have to be used.

1. Finite differences approach

The first approach used to express the acoustic pressure in terms of the discrete strain information is based on a finite differences reconstruction scheme. The displacement field in Eq. (5) is reconstructed from the strain field. This approach makes use of the strain measured in only one direction, say in the x direction.

A finite differences scheme can be used to reconstruct the displacement field $w(x,y)$ from the discrete strain field $\partial^2 w(x,y)/\partial x^2$ known on a $p \times q$ equally spaced grid of points (x_i,y_j) $(i=1,2,...,p;\ j=1,2,...,q)$ on the plate. The central-difference scheme can be written in the following manner:

$$\frac{1}{g_1^2}\begin{bmatrix} -2g_1 & -2 & 2 & & & & \\ & 1 & -2 & 1 & & & \\ & & & \ddots & & & \\ & & & 1 & -2 & 1 & \\ & & & & 2 & -2 & 2g_1 \end{bmatrix}$$

$$\times \begin{bmatrix} w'_{1,1} & w'_{1,2} & \cdots & w'_{1,q} \\ w_{1,1} & w_{1,2} & \cdots & w_{1,q} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ w_{p,1} & w_{p,2} & \cdots & w_{p,q} \\ w'_{p,1} & w'_{p,2} & \cdots & w'_{p,q} \end{bmatrix}$$

$$= \begin{bmatrix} w''_{1,1} & w''_{1,2} & \cdots & w''_{1,q} \\ w''_{2,1} & w''_{2,2} & \cdots & w''_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ w''_{p,1} & w''_{p,2} & \cdots & w''_{p,q} \end{bmatrix}, \tag{9}$$

where $w'_{1,j} \equiv \partial w(x_1,y_j)/\partial x$ and $w'_{p,j} \equiv \partial w(x_p,y_j)/\partial x$ are the slopes of the displacement field in the x direction, on the edges parallel to the y axis, $w''_{k,l} \equiv \partial^2 w(x_k,y_l)/\partial x^2$ is the strain measured in the x direction and g_1 is the spacing between two consecutive measurement points in x. The main interest of the central-difference scheme is that the system is quasitridiagonal and therefore speedy to solve and also economical of memory space to store the coefficients. It should, however, be noted that even if this scheme provides a better approximation to the derivatives than forward or backward differences, its accuracy is limited to the order g_1^2 and a poor accuracy can be expected for large spacing between two consecutive measurement points. The right-hand side of Eq. (9) represents the strain measured at discrete points while the displacement to be reconstructed appears on the left-hand side together with the slopes on $x=x_1$ and $x=x_p$ edges. Equation (9) can be rewritten

$$\mathbf{A}_{p,p+2}\mathbf{w}_{p+2,q} = \mathbf{w}''_{p,q}, \tag{10}$$

where the subscripts indicate the dimensions of the matrices.

It is seen that Eq. (10) can be solved only if two lines of $\mathbf{w}_{p+2,q}$ are known *a priori*. These two lines can be chosen from the first two lines and the last two lines of this matrix, representing boundary conditions on the displacement or the slope along the $x=x_1$ and $x=x_p$ edges. The following equation shows this dependency on boundary conditions:

$$[\mathbf{A}_{p,p}|\mathbf{A}_{p,2}]\begin{bmatrix} \mathbf{w}_{p,q} \\ - \\ \mathbf{w}_{2,q} \end{bmatrix} = \mathbf{w}''_{p,q} \tag{11}$$

or

$$\mathbf{A}_{p,p}\mathbf{w}_{p,q} = \mathbf{w}''_{p,q} - \mathbf{A}_{p,2}\mathbf{w}_{2,q}, \tag{12}$$

where $\mathbf{w}_{2,q}$ represents the two lines of known boundary conditions on the displacement and/or on the slope of the displacement along the edges parallel to the y axis. Equation

1590   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Masson *et al.*: Strain sensing in active control   1590

(12) is solved for $\mathbf{w}_{p,q}$ and the resulting matrix is transformed using a discrete version of Eq. (5) so that the pressure can be calculated by Eq. (2), as detailed later.

In summary, the finite differences approach only requires the measurement of the strain at discrete points on the plate, along the x axis. Two edge conditions on displacement and/or on slope in the x direction have to be specified for the edges parallel to the y axis.

2. Direct approach

The second approach is based on a double integration by parts of the wave-number transform, Eq. (5). The first integration by parts with respect to the x variable leads to the following result for the wave-number transform:

$$\tilde{w}(k_x,k_y)=\frac{1}{ik_x}\oint w(x,y)e^{i(k_xx+k_yy)}\mathbf{n}\cdot\mathbf{i}\,dl$$
$$-\frac{1}{ik_x}\int_{-a}^{a}\int_{-b}^{b}\frac{\partial w(x,y)}{\partial x}e^{i(k_xx+k_yy)}\,dy\,dx$$

$$(13)$$

and the second integration by parts with respect to the x variable gives

$$\tilde{w}(k_x,k_y)=\frac{1}{ik_x}\oint w(x,y)e^{i(k_xx+k_yy)}\mathbf{n}\cdot\mathbf{i}\,dl$$
$$+\frac{1}{k_x^2}\oint\frac{\partial w(x,y)}{\partial x}e^{i(k_xx+k_yy)}\mathbf{n}\cdot\mathbf{i}\,dl$$
$$-\frac{1}{k_x^2}\tilde{S}(k_x,k_y)\qquad(14)$$

with

$$\tilde{S}(k_x,k_y)=\int_{-a}^{a}\int_{-b}^{b}\frac{\partial^2 w(x,y)}{\partial x^2}e^{i(k_xx+k_yy)}\,dy\,dx\qquad(15)$$

and where $\mathbf{i}$ is the unit vector in the x direction, $\mathbf{n}$ is the unit outer vector normal to an edge, $\oint$ denotes the integration over the contour (edges) of the plate, and $\tilde{S}(k_x,k_y)$ is the wave-number transform of the strain field.

The Eq. (14) can be rewritten as

$$\tilde{w}(k_x,k_y)=\sum_{i=1}^{2}I_i-\frac{1}{k_x^2}\tilde{S}(k_x,k_y),\qquad(16)$$

where I_i are the subsisting boundary terms associated with edges $i=1,2$ parallel to the y axis. For a clamped edge (C) parallel to the y axis, where $w(x,y)=\partial w(x,y)/\partial x=0$, the boundary term is simply

$$I_i^C=0.\qquad(17)$$

For a simply supported edge (SS), where $w(x,y)=0$, further integrating by parts the remaining term in Eq. (14) with respect to the y variable leads to the following result:
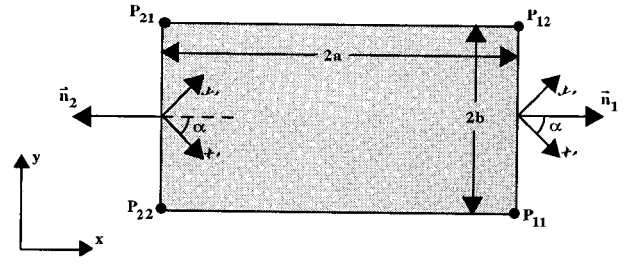
FIG. 3. Change of coordinate system for the elimination of cross derivatives in the direct approach.

$$I_i^{SS}=\frac{1}{ik_x^2k_y}\frac{\partial w(x,y)}{\partial x}e^{i(k_xx+k_yy)}\Big|_{P_{i1}}^{P_{i2}}$$
$$-\frac{1}{ik_x^2k_y}\int_{P_{i1}}^{P_{i2}}\left[\frac{\partial^2 w(x,y)}{\partial x\partial y}\right]_{x=x_i}e^{i(k_xx_i+k_yy)}dy,\quad(18)$$

where P_{i1} and P_{i2} are the end points of the edge i (see Fig. 3). For a free (F) edge, the following result is obtained:

$$I_i^F=I_i^{SS}-\frac{1}{k_xk_y}w(x,y)e^{i(k_xx+k_yy)}\Big|_{P_{i1}}^{P_{i2}}$$
$$+\frac{1}{ik_xk_y^2}\frac{\partial w(x,y)}{\partial y}e^{i(k_xx+k_yy)}\Big|_{P_{i1}}^{P_{i2}}$$
$$-\frac{1}{ik_xk_y^2}\int_{P_{i1}}^{P_{i2}}\left[\frac{\partial^2 w(x,y)}{\partial y^2}\right]_{x=x_i}e^{i(k_xx_i+k_yy)}dy.\quad(19)$$

It appears that for small values of k_x or k_y, divergence may occur in Eqs. (16), (18), or (19). As presented later, three cases will have to be distinguished.

The expressions for I_i^{SS} and I_i^F make boundary terms appear in the calculation of the displacement wave-number spectrum. Boundary conditions at the end points and along the edges parallel to the y axis must then be supplied on the displacement and on its first x and y derivatives (slopes). While no boundary conditions need to be supplied for a clamped (C) edge, the slope in the x direction at the end points together with the cross derivative along the edges parallel to the y axis are requested for a simply supported edge. In addition, the displacement and the slope in the y direction at the end points together with the strain in the y direction along the edges parallel to the y axis are all requested for a free edge. The cross derivatives appearing in the preceding equations are of course not desirable. Fortunately, by using an appropriate change of variable on the boundaries parallel to the y axis as shown in Fig. 3, the cross derivatives (and, as a consequence, the second derivative in y) can be written in terms of the strain measured in the new coordinate system. First, the change of variable on the boundaries is written:

$$
\begin{bmatrix} \dfrac{\partial^2 w(x,y)}{\partial x^2} \\[2mm] \dfrac{\partial^2 w(x,y)}{\partial y^2} \\[2mm] \dfrac{\partial^2 w(x,y)}{\partial x \partial y} \end{bmatrix}
$$

$$
= \begin{bmatrix} \cos^2 \alpha & \sin^2 \alpha & 2\sin\alpha\cos\alpha \\ \sin^2 \alpha & \cos^2 \alpha & -2\sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha & \sin\alpha\cos\alpha & \cos^2\alpha - \sin^2\alpha \end{bmatrix}
$$

$$
\times \begin{bmatrix} \dfrac{\partial^2 w(x,y)}{\partial x'^2} \\[2mm] \dfrac{\partial^2 w(x,y)}{\partial y'^2} \\[2mm] \dfrac{\partial^2 w(x,y)}{\partial x' \partial y'} \end{bmatrix}, \tag{20}
$$

where α is the angle by which the new coordinate system is rotated clockwise with respect to the original one. In order to obtain a simple formulation where all the boundary conditions on the second derivatives of the displacement are expressed using Eq. (20), the cross-derivative in the new coordinate system appearing in this equation is rewritten. First, notice that the bending moment around the direction tangent to the edge, M_n, applied to a simply supported or free edge i is zero:[29,30]

$$
M_n = \begin{bmatrix} n_{ix'} & n_{iy'} \end{bmatrix} \begin{bmatrix} M_{x'} & M_{x'y'} \\ M_{x'y'} & M_{y'} \end{bmatrix} \begin{bmatrix} n_{ix'} \\ n_{iy'} \end{bmatrix} = 0 \tag{21}
$$

with the standard definitions for the internal bending moments under isotropic conditions:

$$
M_{x'} = -\frac{Eh^3}{12(1-\nu^2)} \left(\frac{\partial^2 w(x,y)}{\partial x'^2} + \nu \frac{\partial^2 w(x,y)}{\partial y'^2} \right), \tag{22}
$$

$$
M_{y'} = -\frac{Eh^3}{12(1-\nu^2)} \left(\nu \frac{\partial^2 w(x,y)}{\partial x'^2} + \frac{\partial^2 w(x,y)}{\partial y'^2} \right), \tag{23}
$$

$$
M_{x'y'} = -\frac{Eh^3}{12(1+\nu)} \frac{\partial^2 w(x,y)}{\partial x' \partial y'}, \tag{24}
$$

where $n_{ix'}$, $n_{iy'}$ are, respectively, the x and y components of the normal to the edge i in the new coordinate system and E and ν are, respectively, the Young's modulus and the Poisson ratio of the plate. Solving Eq. (21) with Eqs. (22)–(24), the following equation can then be used to express the cross derivative in the new coordinate system:

$$
\frac{\partial^2 w(x,y)}{\partial x' \partial y'} = -\frac{n_{ix'}}{2n_{iy'}(1-\nu)} \left(\frac{\partial^2 w(x,y)}{\partial x'^2} + \nu \frac{\partial^2 w(x,y)}{\partial y'^2} \right)
$$
$$
-\frac{n_{iy'}}{2n_{ix'}(1-\nu)} \left(\nu \frac{\partial^2 w(x,y)}{\partial x'^2} + \frac{\partial^2 w(x,y)}{\partial y'^2} \right). \tag{25}
$$

With the help of Eqs. (20) and (25), all the second derivatives in Eqs. (18) and (19) can thus be expressed in terms of strain measured in the directions of the new coordinate sys-

tem. For the particular case of $\alpha = \pi/4$, Eq. (20) takes the following form:

$$
\begin{bmatrix} \dfrac{\partial^2 w(x,y)}{\partial x^2} \\[2mm] \dfrac{\partial^2 w(x,y)}{\partial y^2} \\[2mm] \dfrac{\partial^2 w(x,y)}{\partial x \partial y} \end{bmatrix} = \begin{bmatrix} -\dfrac{\nu}{1-\nu} & -\dfrac{\nu}{1-\nu} \\[2mm] \dfrac{1}{1-\nu} & \dfrac{1}{1-\nu} \\[2mm] -\dfrac{1}{2} & \dfrac{1}{2} \end{bmatrix} \begin{bmatrix} \dfrac{\partial^2 w(x,y)}{\partial x'^2} \\[2mm] \dfrac{\partial^2 w(x,y)}{\partial y'^2} \end{bmatrix}. \tag{26}
$$

a. Case k_x small and k_y large. For the case of k_x small, Eq. (14) will obviously diverge. A Taylor series expansion has to be used around that point:

$$
\widetilde{w}(k_x,k_y) = \widetilde{w}(0,k_y) + k_x \lim_{\epsilon \to 0} \widetilde{w}'_{k_x}(\epsilon,k_y)
$$
$$
+ \frac{k_x^2}{2} \lim_{\epsilon \to 0} \widetilde{w}''_{k_x}(\epsilon,k_y), \tag{27}
$$

where $\widetilde{w}'_{k_x}$ is the first derivative, with respect to k_x, of Eq. (16) [with the appropriate expression of I_i given by Eqs. (17)–(19) depending on the boundary conditions] and $\widetilde{w}''_{k_x}$ is the second derivative with respect to k_x. The wave-number transform of the displacement for $k_x = 0$ can be written after integrating twice by parts Eq. (5) with $k_x = 0$:

$$
\widetilde{w}(0,k_y) = \int_{-a}^{a} \int_{-b}^{b} w(x,y) e^{ik_y y} \, dy \, dx \tag{28}
$$

$$
= \oint xw(x,y) e^{ik_y y} \mathbf{n} \cdot \mathbf{i} \, dl
$$
$$
- \int_{-a}^{a} \int_{-b}^{b} x \frac{\partial w(x,y)}{\partial x} e^{ik_y y} \, dy \, dx \tag{29}
$$

$$
= \oint xw(x,y) e^{ik_y y} \mathbf{n} \cdot \mathbf{i} \, dl + \int_{-a}^{a} \int_{-b}^{b} \left(\frac{x^2}{2} - \frac{a^2}{2} \right)
$$
$$
\times \frac{\partial^2 w(x,y)}{\partial x^2} e^{ik_y y} \, dy \, dx. \tag{30}
$$

b. Case k_x large and k_y small. In a similar way, a series expansion has to be used in this case:

$$
\widetilde{w}(k_x,k_y) = \widetilde{w}(k_x,0) + k_y \lim_{\delta \to 0} \widetilde{w}'_{k_y}(k_x,\delta)
$$
$$
+ \frac{k_y^2}{2} \lim_{\delta \to 0} \widetilde{w}''_{k_y}(k_x,\delta), \tag{31}
$$

where $\widetilde{w}'_{k_y}$ is the first derivative, with respect to k_y, of Eq. (16) [again with the appropriate expression of I_i given by Eqs. (17)–(19)] and $\widetilde{w}''_{k_y}$ is the second derivative with respect to k_y. The wave-number transform of the displacement for $k_y = 0$ is given directly from Eq. (14) by

$$\widetilde{w}(k_x,0)=\frac{1}{ik_x}\oint w(x,y)e^{ik_x x}\mathbf{n}\cdot\mathbf{i}\,dl$$

$$+\frac{1}{k_x^2}\oint\frac{\partial w(x,y)}{\partial x}e^{ik_x x}\mathbf{n}\cdot\mathbf{i}\,dl$$

$$-\frac{1}{k_x^2}\int_{-a}^{a}\int_{-b}^{b}\frac{\partial^2 w(x,y)}{\partial x^2}e^{ik_x x}\,dy\,dx. \quad (32)$$

c. Case k_x small and k_y small. In this case, the series expansion takes the form:

$$\widetilde{w}(k_x,k_y)=\widetilde{w}(0,0)+k_x\lim_{\substack{\delta\to 0\\\epsilon\to 0}}\widetilde{w}'_{k_x}(\epsilon,\delta)$$

$$+k_y\lim_{\substack{\delta\to 0\\\epsilon\to 0}}\widetilde{w}'_{k_y}(\epsilon,\delta)+\frac{k_x^2}{2}\lim_{\substack{\delta\to 0\\\epsilon\to 0}}\widetilde{w}''_{k_x}(\epsilon,\delta)$$

$$+\frac{k_y^2}{2}\lim_{\substack{\delta\to 0\\\epsilon\to 0}}\widetilde{w}''_{k_y}(\epsilon,\delta)+k_x k_y\lim_{\substack{\delta\to 0\\\epsilon\to 0}}\widetilde{w}''_{k_x k_y}(\epsilon,\delta),$$

$$(33)$$

where $\widetilde{w}'_{k_x k_y}$ is the cross derivative of Eq. (14) with respect to k_x and k_y. The wave-number transform of the displacement for $k_x=0$ and $k_y=0$ takes the form:

$$\widetilde{w}(0,0)=\oint xw(x,y)\mathbf{n}\cdot\mathbf{i}\,dl+\int_{-a}^{a}\int_{-b}^{b}\left(\frac{x^2}{2}-\frac{a^2}{2}\right)$$

$$\times\frac{\partial^2 w(x,y)}{\partial x^2}\,dy\,dx. \quad (34)$$

In summary, this approach requires, in the general case, the measurement of strain along the x axis at discrete points on the plate and, along the x' and y' axes on the boundaries parallel to the y axis. Moreover, except for the case of a clamped edge, end-points conditions have to be specified on the displacement and its first derivatives. Although this extensive required knowledge might appear to be a major drawback, the approach can turn to be very effective if the boundary terms in Eq. (16) can be neglected (to be examined later). In the following, this approach will be referred to as the direct approach.

II. VALIDATION OF THE APPROACHES

A numerical simulation is performed to assess the performance of both the finite differences approach and the direct approach. The setting for this simulation is first described and following that, the acoustic power radiated using both approaches is compared with an analytical solution.[27] The proper validation of the approaches for the calculation of the acoustic power from strain information is necessary for the method to be applied in an active control context.

A. Setting

The setting used for the simulation is shown in Fig. 1, without considering the PZT for the moment. The physical parameters of the plate are presented in Table I. In addition, the plate is assumed to be thin and simply supported and it is

TABLE I. Components characteristics.

	Plate	Actuator(s)
Material	steel	PZT patch (PSI-5A-S2)
x dimension (m)	0.455	0.06
y dimension (m)	0.379	0.04
Young's modulus E (GPa)	210	61
Thickness h (m)	0.003	0.0002
Poisson ratio ν	0.3	0.29
Density ρ (kg/m^3)	7850	—
Structural damping η	0.01	—

excited by a point force disturbance of amplitude 1 N and frequency ω_0, located at $(\hat{x},\hat{y})=(a/2,b/2)=(0.1138,0.0948)$.

B. Disturbance equation

A modal expansion is used to simulate the plate's response to the disturbance. Such an expansion is used for the displacement $W(x,y,t)$ which can be written as

$$W(x,y,t)=\sum_{m,n}a_{mn}\hat{w}_{mn}(x,y)e^{i\omega_0 t};\quad m,n=1,2,..., \quad (35)$$

where the a_{mn} are the modal coefficients, ω_0 is the excitation frequency, and $\hat{w}_{mn}(x,y)$ are the eigenfunctions given, in the case of a simply supported plate, by

$$\hat{w}_{mn}(x,y)=\sin\left(\frac{m\pi(x+a)}{2a}\right)\sin\left(\frac{n\pi(y+b)}{2b}\right). \quad (36)$$

Under a point force disturbance, the modal coefficients are given by

$$a_{mn}=-\frac{F\hat{w}_{mn}(\hat{x},\hat{y})}{\rho hab(\omega_0^2-\omega_{mn}^2-i\eta\omega_{mn}^2)}, \quad (37)$$

where F is the amplitude of the force applied at $(\hat{x},\hat{y})$, ρ is the mass density of the plate, η is the structural damping ratio, and ω_{mn} is the modal frequency given by

$$\omega_{mn}=\sqrt{\frac{Eh^2}{12\rho(1-\nu^2)}}[\gamma_m^2+\gamma_n^2], \quad (38)$$

where

$$\gamma_m=\frac{m\pi}{2a}, \quad (39)$$

$$\gamma_n=\frac{n\pi}{2b}. \quad (40)$$

The modal frequencies of the plate are presented in Table II.

C. Finite differences approach

Within this approach, the strain field is simulated at discrete points on the structure (see Fig. 2), thereby simulating sensor measurements, and provided as the input matrix $\mathbf{w}''_{p,q}$ in Eq. (12). In the case of a simply supported plate, the matrices $\mathbf{w}_{p,q}$ and $\mathbf{w}_{2,q}$ reduce to:

TABLE II. Plate modal frequencies.

Order	Mode	Modal frequency (Hz)
1	(1,1)	87
2	(2,1)	194
3	(1,2)	241
4	(2,2)	348
5	(3,1)	372
6	(1,3)	497
7	(3,2)	526
8	(2,3)	604
9	(4,1)	621
10	(4,2)	775
11	(3,3)	782

$$\mathbf{w}_{p,q} = \begin{bmatrix} w'_{1,1} & w'_{1,2} & \cdots & w'_{1,q} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ w_{p-1,1} & w_{p-1,2} & \cdots & w_{p-1,q} \\ w'_{p,1} & w'_{p,2} & \cdots & w'_{p,q} \end{bmatrix}, \qquad (41)$$

$$\mathbf{w}_{2,q} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,q} \\ w_{p,1} & w_{p,2} & \cdots & w_{p,q} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \qquad (42)$$

The matrix $\mathbf{w}_{p,q}$, obtained by solving Eq. (12), is then transformed using a discretized version of Eq. (5), substituted in Eq. (2) and in Eq. (7) to obtain the radiated acoustic power.

The effect of the number of sensors is first studied for this approach. The wave-number transform [Eq. (5)] is evaluated as a discrete wave-number transform, similar to a discrete Fourier transform:

$$\widetilde{w}(\lambda,\mu) = g_1 g_2 \sum_{i=1}^{p} \sum_{j=1}^{q} w_{i,j} e^{i(\lambda x_i + \mu y_j)}, \qquad (43)$$

where g_1 and g_2 are the spacing between measurement points, respectively, in x and y directions. By using this equation, the integral in Eq. (7) is made over continuous wave numbers using an adaptive two-dimensional quadrature.

Figure 4 presents a comparison for the finite differences approach using a 6×6 sensing grid and a 16×16 sensing grid. It is seen that with a sufficient number of sensors, the agreement with the analytical solution is very good over the whole range. When the number of sensors is 6×6, good agreement is obtained up to 500 Hz. It appears that, excluding the sensors located on the boundaries, at least two sensors per mechanical wavelength are required in a given direction to obtain adequate estimate of the radiated acoustic power.

In order for it to be applied to active control, it is desirable to improve the time-domain performance of the approach. In this respect, the wave-number transform can be calculated faster using a fast Fourier transform (FFT) algorithm, which means that the integral in Eq. (7) is now made over discrete wave numbers. Considering the modal expansion used in the following for a simply supported beam [Eq.
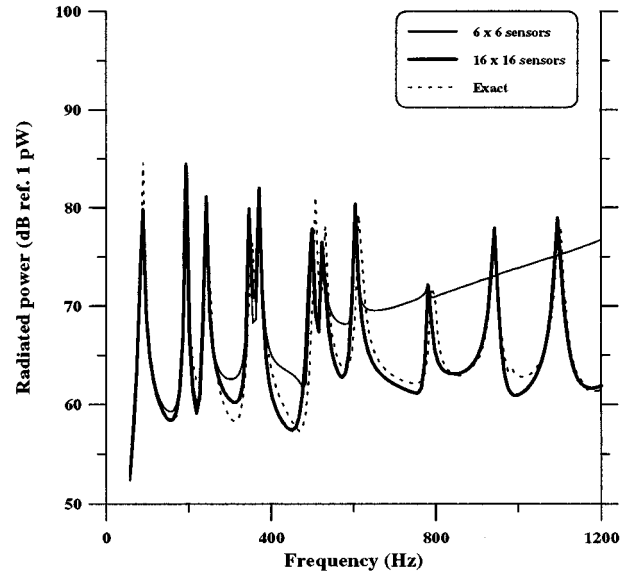
FIG. 4. Effect of the number of sensors with continuous wave-number transform—finite differences approach.

(35)], the wave-number spectrum of the displacement is either symmetric or antisymmetric and, therefore, the analysis can be restricted to positive wave numbers only. The transform is expressed as

$$\widetilde{w}(\lambda_u,\mu_v) = g_1 g_2 \sum_{i=1}^{p} \sum_{j=1}^{q} w_{i,j} e^{i(\lambda_u x_i + \mu_v y_j)} \qquad (44)$$

with the discretized wave numbers:

$$\lambda_u = 2\pi \frac{u}{\hat{p} g_1}, \quad u = 0,...,\hat{p}-1, \qquad (45)$$

$$\mu_v = 2\pi \frac{v}{\hat{q} g_2}, \quad v = 0,...,\hat{q}-1, \qquad (46)$$

where $\hat{p} \geq p$ and $\hat{q} \geq q$ are the number of points to be used in the FFT calculation. The radiated acoustic power is then calculated using a discretized version of Eq. (7):

$$\Pi_d = \frac{\omega^3 \rho_0}{2 \hat{p} \hat{q} g_1 g_2} \sum_{u=0}^{\hat{p}-1} \sum_{v=0}^{\hat{q}-1} \text{Re} \left[\frac{|\widetilde{w}(\lambda_u,\mu_v)|^2}{\sqrt{k^2 - \lambda^2 - \mu^2}} \right]. \qquad (47)$$

The implementation of this algorithm is not straightforward as the resolution required on the components in the summation of Eq. (47) gets more and more important as the frequency decreases. A simple way to increase the accuracy is to use more interpolation points by using enough zero-padding. By doing so, $\hat{p}$ and $\hat{q}$ are increased and more points are considered for λ_u and μ_v. Figure 5 shows the effect of the number of zeros added on each of the four sides of a 16×16 grid sensing. As expected, the major effect of increasing the number of zeros is seen at low frequencies where only few wave-number components are used and where, therefore, the number of the components considered in Eq. (47) as well as the precision on their calculation is crucial.
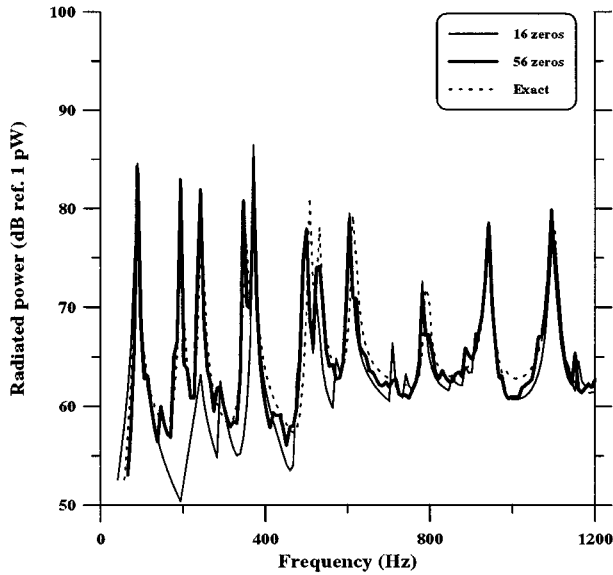
FIG. 5. Effect of zero-padding with discrete wave-number transform—finite differences approach with 16×16 sensors.

FIG. 6. Singularity treatment in the direct approach, $k_y = 19.5$.

D. Direct approach

In the case for which λ and μ are large, the I_i terms appearing in Eq. (16) take, for a simply supported plate and after discretization, the following form:

$$I_i = -\frac{1}{i\lambda^2\mu} \, g_2 \sum_{j=1}^{q} \left[\frac{\partial^2 w(x,y)}{\partial x \partial y} \right]_{\substack{x=x_i \\ y=y_j}} e^{i(\lambda x_i + \mu y_j)}. \tag{48}$$

Similar expressions are obtained for small values of λ and μ.

The strain field is simulated at discrete sensors points, wave number transformed over continuous wave numbers and, together with the simulated I_i given by Eq. (48), substituted into Eqs. (16), (2), and (7) to obtain the radiated acoustic power. Following the λ and μ values in the integral of Eq. (7), either Eq. (27), (31), or (33) may have to be used instead of Eq. (16). The comparison of Eqs. (16) and (27) is shown in Fig. 6 for a given value of k_y. It is seen that Eq. (27) is an appropriate approximation for k_x small. The cross derivative in Eq. (48) is obtained by simulating strain measurements along the x' and y' axes by Eqs. (20) and (25).

Figure 7 presents a comparison for the direct approach using a 6×6 sensing grid and a 16×16 sensing grid. It appears that this approach offers good agreement with the analytical reference for a sufficient number of sensors but that this agreement deteriorates rapidly with a diminution in the number of sensors.

The wave-number transform can also be implemented using an FFT algorithm in this approach. The sensibility to zero-padding is similar to the one observed with the preceding approach.

As shown in Figs. 4 and 7, the finite differences approach seems to have a lower requirement on the number of sensors in order to agree with an analytical reference. Moreover, the formulation of the first approach is relatively simple and well-suited for computer coding in real-time applications such as active control.

Using all the terms in the direct approach makes it relatively complicated, compared to the finite differences approach. It will be shown, however, that at least for a simply supported plate, good control performance can be obtained at low frequency using the direct approach with the boundary terms neglected.

III. OPTIMAL CONTROL RESULTS

A. Setting

In addition to the setting presented in Sec. II A, Table I presents the physical parameters of the PZT actuators used for the control. The piezoelectric strain constant is $d_{31} = 1.71^{-10}$ m/V. Control results are presented for a spanning
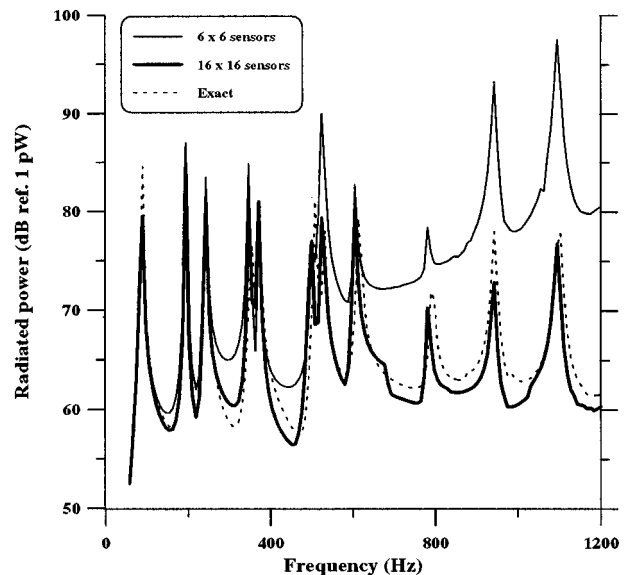
FIG. 7. Effect of the number of sensors with continuous wave-number transform—direct approach.

of the excitation frequency between 0 and 1.2 kHz. The control is then performed at a single frequency using two actuators configurations. The first one consists of one PZT patch centered at $(x,y)=(-0.08,0.07)$ and the second configuration consists of four PZT patches centered symetrically with respect to the axes at $(x,y)=(-0.08,0.07)$; $(-0.08, -0.07)$; $(0.08,-0.07)$; $(0.08,0.07)$.

In the context of active control, the resulting pressure at a point M is the sum of a primary pressure field (labeled p) created by a disturbance, i.e., a point force, and secondary pressure fields (labeled s) originating from n actuators exciting the structure. The total displacement field of the plate is thus given by

$$w(x,y)=w_p(x,y)+\mathbf{w}_s(x,y)\cdot\mathbf{V}_s, \qquad (49)$$

where w_p is the primary displacement, $\mathbf{w}_s$ is an n-dimensional vector defined as

$$\mathbf{w}_s(x,y)=[\cdots \quad w_s^i(x,y) \quad \cdots], \qquad (50)$$

where the ith element represents the displacement field resulting from the ith actuator ($i=1,...,n$) and $\mathbf{V}_s$ is the dimension n vector of the control variable, to be adjusted so as to minimize the radiated power, and defined as

$$\mathbf{V}_s=\begin{bmatrix} \vdots \\ V_s^i \\ \vdots \end{bmatrix}. \qquad (51)$$

Considering Eqs. (2), (16), and (44), the discretized pressures are obtained from

$$p_{uv}=-\frac{\omega^2\rho_0}{2\pi}\frac{e^{-ikR}}{R} g_1 g_2 \sum_{i=1}^{p}\sum_{j=1}^{q} w_{i,j}e^{i(\lambda_u x_i + \mu_v y_j)} \quad (52)$$

for the finite differences approach and by

$$p_{uv}=-\frac{\omega^2\rho_0}{2\pi}\frac{e^{-ikR}}{R}\left[\sum_{i=1}^{2} I_i\right.$$
$$\left.-\frac{1}{\lambda_u^2} g_1 g_2 \sum_{k=1}^{p}\sum_{j=1}^{q} w_{k,j}''e^{i(\lambda_u x_k + \mu_v y_j)}\right] \qquad (53)$$

for the direct approach. This discretized knowledge of the pressure means that sufficient zero-padding is required for adequate calculation of the summation in Eq. (47).

B. Actuators equations

A modal expansion is also used to simulate the plate's response to the actuators [see Eq. (35)]. For a piezoelectric patch (PZT) extending from x_1 to x_2 and from y_1 to y_2, the modal coefficients are calculated as follows[6] for a simply supported plate:

$$a_{mn}=\frac{C_0\epsilon_{pe}}{\rho hab(\omega_0^2-\omega_{mn}^2-i\eta\omega_{mn}^2)}\left[\frac{\gamma_m^2+\gamma_n^2}{\gamma_m\gamma_n}(\cos\gamma_m(x_1+a)-\cos\gamma_m(x_2+a))(\cos\gamma_n(y_1+b)-\cos\gamma_n(y_2+b))\right], \qquad (54)$$

where
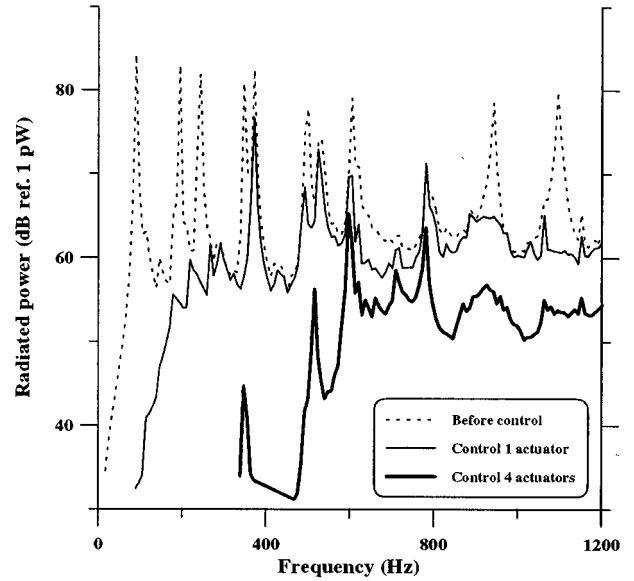
FIG. 8. Active control using 16×16 strain sensors and discrete wave-number transform—finite differences approach.

$$C_0=-E\left[\frac{1+\nu_{pe}}{1-\nu}\right]\frac{P}{1+\nu-(1+\nu_{pe})P}\frac{h^2}{6}, \qquad (55)$$

$$P=-\frac{E_{pe}}{E}\left[\frac{1-\nu^2}{1-\nu_{pe}^2}\right]K, \qquad (56)$$

$$K=\frac{3hh_{pe}(h+h_{pe})}{4(h^3/8+h_{pe}^3)+(3/2)hh_{pe}^2}, \qquad (57)$$

$$\epsilon_{pe}=\frac{d_{31}}{h_{pe}}V, \qquad (58)$$

and where ν_{pe}, E_{pe}, and h_{pe} are the Poisson ratio, Young's modulus, and thickness of the piezoelectric and V is the voltage applied to the PZT. This voltage is the control variable and corresponds to an element of $\mathbf{V}_s$.

C. Finite differences approach

The primary and secondary strain fields are simulated on a discrete $p\times q$ grid of points equally spaced on the structure, thereby simulating sensors measurements, and each of these fields is successively provided as the matrix $\mathbf{w}_{p,q}''$ in Eq. (12). The matrix $\mathbf{w}_{p,q}$ obtained by solving Eq. (12) is then substituted in Eqs. (49) and (52) so that a discretized form of Eq. (7) similar to Eq. (47) is obtained which is a quadratic form to be minimized by calculating the optimal control vector $\mathbf{V}_s$.

Results obtained under the two actuators configurations are presented in Fig. 8 for a 16×16 sensing grid (with 56 zeros on each side). It is seen that increasing the number of actuators has a dramatic effect on the control performance, especially at low frequencies. However, few peaks are poorly controlled due to the particular location of the actuators as they cannot excite those modes efficiently. No optimization was performed on the location of the actuators.
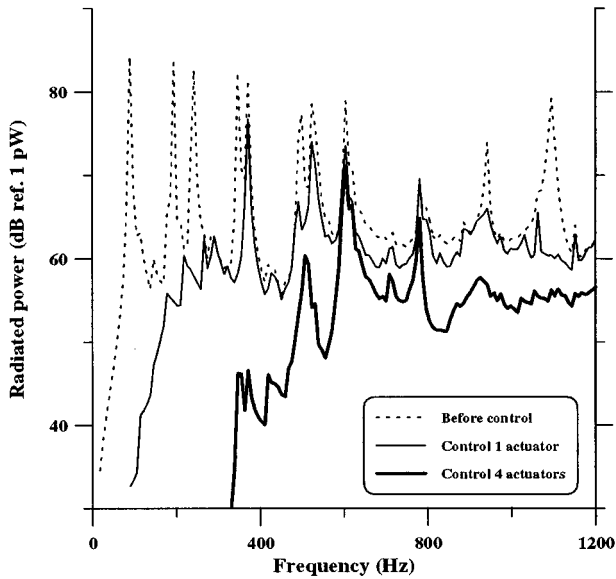
FIG. 9. Active control using 16×16 strain sensors and discrete wave-number transform—direct approach.

FIG. 10. Effect of the boundary terms on active control using the direct approach and discrete wave-number transform.

D. Direct approach

Within the direct approach, the strain field is also simulated at discrete sensors points and, together with Eq. (49), substituted in the simulated I_i given by Eq. (48) and in Eq. (53). A quadratic form based on a discrete version of Eq. (7) is then minimized to obtain the optimal control vector $\mathbf{V}_s$.

Figure 9 presents optimal control results for a 16×16 sensing grid. In this situation, as with the preceding approach, a major improvement is observed in the control performance when the number of actuators goes from one to four.

This approach could become very interesting if the boundary terms I_i can be neglected in the cost function, i.e., sensing strain only within plate surface and using the wave-number transform of the strain as a cost function. It was observed that, for a given mode, the boundary transform term [Eq. (18)] and the surface strain transform term [Eq. (15)] are of the same order of magnitude, below the modal frequency. The importance of the surface strain transform term increases slightly close to that frequency while its relative importance decreases rapidly above this frequency. The global effect is that, with an increase in frequency, more and more modes are considered in which the surface strain transform term has little effect or, conversely, the boundary term has large relative effect. Consequently, it is expected that good optimal control performance could be obtained at low frequencies. In fact, Fig. 10 shows that, with 16×16 sensors and four actuators, rather good control is achieved when this approach is used without the boundary terms I_i not only at low frequencies but also at higher frequencies, thereby allowing for good control performance over a wide band at low frequencies. This result also illustrates the low sensitivity of the approach to correct boundary conditions used in Eq. (48).
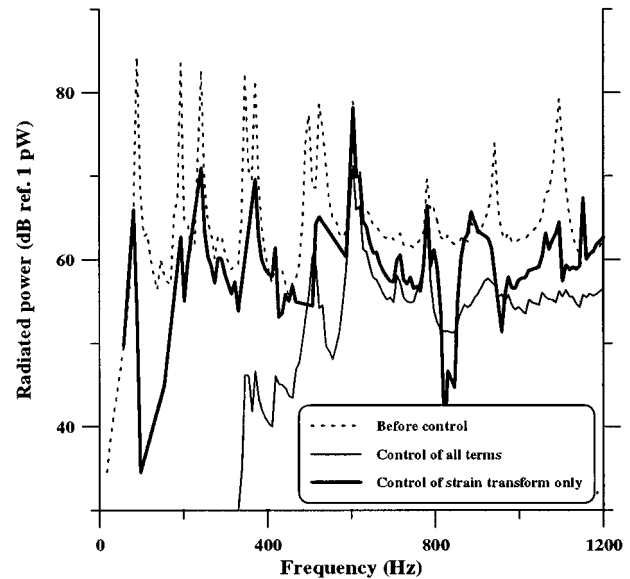
E. Comparison of the approaches

Both approaches already offer very interesting control performance with few sensors and with only one actuator. However, in this situation, as the frequency increases to 1 kHz, the direct approach seems to offer better performance. An increase in the number of sensors tends to eliminate this discrepancy. Then, the finite differences approach presents slightly better reduction in the low-frequency range with four actuators. An increase in the number of actuators is more beneficial to the finite differences approach than to the direct approach. It should be noted that, due to the known boundary conditions in this case, the number of sensors is effectively reduced by two in each direction.

As they involve operations such as FFT, matrix inversions and numerical integration, both approaches seem to be computationally intensive. Thus in order to use them in real-time adaptive control strategy, the control algorithm could be divided into two parts:[31] the control part which executes in real time and the optimization part which is calculated during the idle processor time using a block processing technique.[32] With this approach, higher sampling frequencies could be sustained for a given computational power. This strategy implies, however, slower adaptation to time varying statistics of the signal.

F. Effect on wave-number spectrum

To better illustrate the effect of the control on the plate's behavior, the wave-number spectra are presented before and after control with a 16×16 sensing grid and four actuators to perform the control.

The excitation frequency is chosen to be located on a resonance [mode (1,3), 497 Hz] where the control has an important effect. The set of Figs. 11 and 12 thus presents the effect of control in the on-resonance case. As it was noticed before,[8] the acoustic power reduction is achieved through the
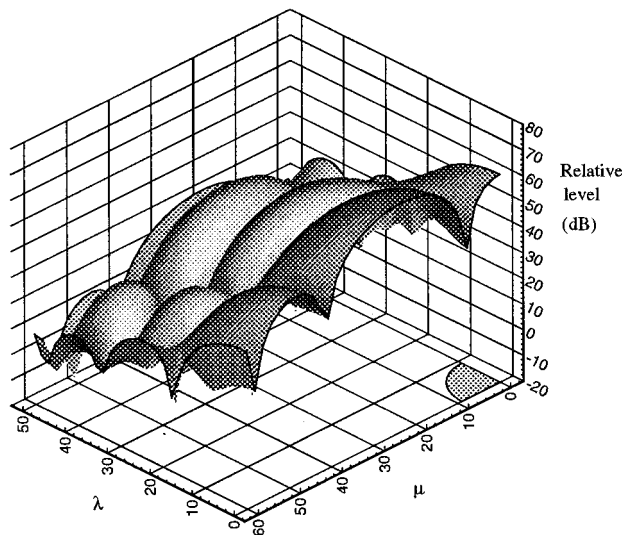
FIG. 11. Wave-number spectrum of the displacement, before control, on-resonance (497 Hz).

suppression of the supersonic wave-number components, represented in a shaded area close to the origin of the (λ, μ) plane.

IV. CONCLUSIONS

A new control strategy was presented in which the cost function, defined as the radiated acoustic power, involves monitoring the strain field at discrete points on the structure. Two approaches were proposed in order to use this discrete strain information directly in the cost function. Being defined in the wave-number domain, this cost function is minimized with a concentration of the control effort in the supersonic (radiating) region of the wave-number spectrum of the displacement. The first approach presented is based on a recon-

struction of the displacement field from the strain field while the second one is based on the integration by parts of the wave-number transform of the displacement.
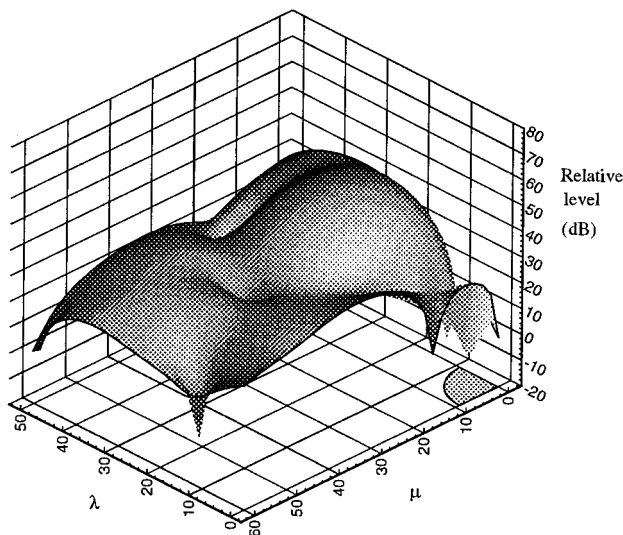
Both approaches were validated by comparison with an analytical reference for the calculation of the radiated acoustic power (cost function) from a simply supported plate. Good performance was obtained with at least two strain sensors per structural wavelength. The finite differences approach appeared to be simple and effective in representing the radiated acoustic power from a simply supported plate. Although the direct approach may suffer from extensive knowledge required on the boundary conditions, it was shown that, at least for the simply supported plate, the effect of these boundary terms is weak at low frequencies and increasing with the frequency and, therefore, this approach seems appropriate for active control at low frequencies.

Optimal control results obtained using both approaches have demonstrated that the radiated acoustic power from a simply supported plate can be reduced significantly over a wide frequency band by using 16×16 discrete strain sensors on the plate (which in fact reduce to 14×14 sensors on the plate due to the knowledge of the boundary conditions) and only one actuator. An increase in the number of actuators produce even more reduction. Under these conditions, the finite differences approach seems to perform better. Control performance of the direct approach, without the boundary terms, using a 16×16 array of sensors and four actuators is seen to be interesting, considering the simplicity of the approach in this case. It has been shown that the supersonic region of the displacement spectrum is effectively reduced and that the control effort is thus spent reducing amplitudes in this region.

In its simplest form, the proposed ASAC strategy is easy to implement and is very promising for application in real control situations over a wide frequency band and using direct strain measurements. Implementation of this strategy into standard adaptive control algorithms is currently being investigated. Future work also includes validation for plates with different support conditions.

FIG. 12. Wave-number spectrum of the displacement, control with four actuators, on-resonance (497 Hz).

[1] S. J. Elliot and P. A. Nelson, ''Active noise control,'' IEEE Signal Process. Mag. **10**, 12–35 (1993).
[2] C. R. Fuller, ''Active control of sound transmission/radiation from elastic plates by vibration inputs: I. Analysis,'' J. Sound Vib. **136**, 1–15 (1990).
[3] S. J. Elliott, ''Active control of structure-borne noise,'' J. Sound Vib. **177**, 651–673 (1994).
[4] M. V. Gandhi and B. S. Thompson, *Smart Materials and Structures* (Chapman and Hall, London, 1992).
[5] E. F. Crawley, ''Use of piezoelectric actuators as elements of intelligent structures,'' AIAA J. **25**, 1373–1385 (1987).
[6] E. K. Dimitriadis, ''Piezoelectric actuators for distributed vibration excitation of thin plates,'' ASME J. Vib. Acoust. **113**, 100–107 (1991).
[7] C. R. Fuller, C. H. Hansen, and S. D. Snyder, ''Experiments on active control of sound radiation from a panel using a piezoceramic actuator,'' J. Sound Vib. **150**, 179–190 (1991).
[8] Y. Gu and C. R. Fuller, ''Active control of sound radiation from a fluid-

loaded rectangular uniform plate,'' J. Acoust. Soc. Am. **93**, 337–345 (1993).

[9] J. D. Maynard, E. G. Williams, and Y. Lee, ''Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH,'' J. Acoust. Soc. Am. **78**, 1385–1413 (1985).

[10] W. T. Baumann, W. R. Saunders, and H. H. Robertshaw, ''Active structural acoustic control of broadband disturbances,'' J. Acoust. Soc. Am. **92**, 1998–2005 (1992).

[11] J. P. Maillard and C. R. Fuller, ''Advanced time domain wave-number sensing for structural acoustic systems. Part I. Theory and design,'' J. Acoust. Soc. Am. **95**, 3252–3261 (1994).

[12] J. P. Maillard and C. R. Fuller, ''Advanced time domain wave-number sensing for structural acoustic systems. Part II. Active radiation control of a simply supported beam,'' J. Acoust. Soc. Am. **95**, 3262–3272 (1994).

[13] J. P. Maillard and C. R. Fuller, ''Advanced time domain wave-number sensing for structural acoustic systems. Part III. Experiments on active broadband radiation control of a simply supported plate,'' J. Acoust. Soc. Am. **98**, 2613–2621 (1995).

[14] R. L. Clark and C. R. Fuller, ''Modal sensing of efficient acoustic radiators with polyvinylidene fluoride distributed sensors in active structural acoustic control approaches,'' J. Acoust. Soc. Am. **91**, 3321–3329 (1992).

[15] C. K. Lee and F. C. Moon, ''Modal sensors/Actuators,'' ASME J. Appl. Mech. **57**, 434–441 (1990).

[16] S. D. Snyder, C. H. Hansen, and N. Tanaka, ''Shaped vibration sensors for feedforward control of structural radiation,'' in *Proceedings of the Second Conference on Recent Advances in Active Control of Sound and Vibration* (Technomic, 1993), pp. 177–188.

[17] M. E. Johnson and S. J. Elliott, ''Active control of sound radiation using volume velocity cancellation,'' J. Acoust. Soc. Am. **98**, 2174–2186 (1995).

[18] M. E. Johnson and S. J. Elliott, ''Experiments on the active control of sound radiation using a volume velocity sensor,'' in *Proceedings of the SPIE North American Conference on Smart Structures and Materials* (1995), pp. 658–669.

[19] C. Guigou, A. Berry, F. Charette, and J. Nicolas, ''Active control of finite beam volume velocity using shaped PVDF sensor,'' Acta Acust. (China) **82**, 772–783 (1996).

[20] F. Charette, C. Guigou, and A. Berry, ''Development of volume velocity sensors for plates using PVDF film,'' in *Proceedings of Active 95*, edited by S. Sommerfeldt and H. Hamada (Noise Control Foundation, 1995), pp. 241–252.

[21] S. D. Sommerfeldt and B. L. Scott, ''Estimating acoustic radiation using wavenumber sensors,'' in *Proceedings of Noise-Con 94* (Noise Control Foundation, 1994), pp. 279–284.

[22] R. L. Clark and C. R. Fuller, ''A model reference approach for implementing active structural acoustic control,'' J. Acoust. Soc. Am. **92**, 1534–1544 (1992).

[23] W. B. Spillman and B. R. Kline, ''Fiber optic vibration sensors for structural control applications,'' in *Proc. Damping 89 Conf.* (1989), pp. ICA-1–ICA-21.

[24] R. L. Clark, C. R. Fuller, B. R. Fogg, W. V. Miller, A. M. Vengsarkar, and R. O. Claus, ''Active structural acoustic control using fiber sensors and piezoelectric actuators,'' in *Proceedings of the ADPA/AIAA/ASME/SPIE Conference on Active Materials and Adaptive Structures* (IOP, 1992), pp. 547–552.

[25] C. R. Fuller and R. A. Burdisso, ''A wavenumber domain approach to the active control of structure-borne sound,'' J. Sound Vib. **148**, 355–360 (1991).

[26] S. D. Sommerfeldt, ''Active wavenumber control of acoustic radiation from a plate,'' in *Proceedings of the Second Conference on Recent Advances in Active Control of Sound and Vibration* (Technomic, 1993), pp. 929–940.

[27] A. Berry, J. L. Guyader, and J. Nicolas, ''A general formulation for the sound radiation from rectangular, baffled plates with arbitrary boundary conditions,'' J. Acoust. Soc. Am. **88**, 2792–2802 (1990).

[28] R. L. Clark and C. R. Fuller, ''Active structural acoustic control with adaptive structures including wavenumber considerations,'' in *Proceedings of the Conference on Recent Advances in Active Control of Sound and Vibration* (Technomic, 1991), pp. 507–524.

[29] L. Meirovitch, *Analytical Methods in Vibrations* (MacMillan, London, 1967).

[30] J.-L. Batoz and G. Dhatt, *Modélisation des structures par éléments finis. 2. Poutres et plaques* (Hermès, Paris, 1990).

[31] M. Bouchard and B. Paillard, ''A transform domain optimization to increase the convergence speed of the multichannel filtered-*X* least-mean-square algorithm,'' J. Acoust. Soc. Am. **100**, 3203–3214 (1996).

[32] S. Haykin, *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1996), 3rd ed.

Vibration analysis of symmetrically laminated thick rectangular plates using the higher-order theory and p-Ritz method

C. C. Chen
Department of Civil Engineering, The University of Queensland, Brisbane, Queensland 4072, Australia

K. M. Liew
Division of Engineering Mechanics, School of Mechanical and Production Engineering,
Nanyang Technological University, Singapore 639798

C. W. Lim and S. Kitipornchai
Department of Civil Engineering, The University of Queensland, Brisbane, Queensland 4072, Australia

The free vibration analysis of symmetrically laminated thick rectangular plates is examined. The p-Ritz method is employed in which sets of uniquely defined polynomials are used as the admissible trial displacement and rotation functions. The energy integral expressions of the laminates are derived by incorporating the shear deformation using Reddy's higher-order plate theory [J. Appl. Mech. ASME **51**, 745–752 (1984)]. The formulation is basically applicable to rectangular laminates with any combination of free, simply supported, and clamped boundary conditions. To evaluate the validity and to demonstrate the applicability of the proposed method, a series of free vibration analyses of laminated composite plates is reported. Wherever possible, the accuracy of this analysis is validated through comparison with available results. Efforts are made to interpret the results to provide physical insight to the problem. © *1997 Acoustical Society of America.*
[S0001-4966(97)01708-6]

PACS numbers: 43.40.Dx [CBB]

INTRODUCTION

The unique properties of composite laminates, including high strength and high stiffness-to-weight ratios, high fatigue resistance, high damping, and potential for tailoring, have led to serious examination of their engineering performance in areas such as vibration, buckling, and stress analyses. An understanding of the vibration behavior of composite panels has particularly attracted many researchers to the possibility of furnishing an optimal design state. Numerous works have applied a direct extension of Kirchhoff's classical theory of plates to composite laminates. The assumptions underlying this theory, however, result in underestimation of deflection and overestimation of natural frequencies of a plate. Because of this disadvantage in the classical plate theory, numerous refined theories incorporating the transverse shear deformations have been proposed.

In existing literature, the first-order laminated plate theory was due to Yang *et al.*,[1] which was extended to the higher-order laminated plate theory by Reddy.[2] The first-order theory implies a conceptual paradox as the transverse shear strain does not vanish on the top and bottom surfaces. The thick laminated plate theories were developed to improve the modeling of transverse shear distribution. In addition, the higher-order theory discards the shear correction factors required in the first-order theory. A comprehensive review of the various refined plate theories for laminates proposed over the years has been summarized by Reddy and Robbins Jr.[3] and Noor and Burton.[4] Out of these theories, the theory used in the present study comes under the class of a single-layer displacement-based theory.[2] In this theory, the three-dimensional elasticity theory is reduced to two dimensions by replacing the laminated plate with an equivalent homogenous anisotropic plate and introducing a global displacement approximation in the thickness direction. The order of approximation is with respect to the distribution of displacements in the thickness direction. Thus, this theory provides reasonably accurate solutions in predicting the global behavior of composite laminates.

During the past decades, there has been much interest in the finite element method using various laminate plate theories because of its versatility.[5–7] However, sufficient numbers of discretized elements and nodes are required if there are curved boundary conditions. The p-Ritz method[8] somewhat improves these shortcomings by using the more traditional Rayleigh–Ritz method which assumes the entire plate as a single element and eliminates the need for discretization, mesh generation, and larger degrees of freedom. Excellent results were seen on earlier works for isotropic plates and thin and thick laminates.[9–12] An important feature of the Ritz method is the selection of admissible functions in the series representing the unknown functions in the displacement field. The accuracy and convergence of the solution are greatly dependent on the choice of the trial functions. In the present analysis, the p-Ritz functions are employed in which sets of uniquely defined polynomials are used as the admissible trial displacement and rotation functions. To account for the transverse shear effects, Reddy's higher-order shear deformation plate theory has been integrated with the p-Ritz method for solving the free vibration analysis of thick laminated plates with various boundary conditions.

It should be noted that Reddy and his associates[2,5] have presented results for only plates with opposite sides simply

supported. Results for plates with general boundary conditions are still unavailable, to the authors' knowledge. Therefore, this paper attempts to provide examples of rectangular plates with various boundary conditions to show the applicability and versatility of the p-Ritz method, without the difficulty of mesh generation and continuity conditions of other discretization methods. Details of the analytical method and formulation for the problem of free vibration of laminated plates with combinations of clamped, simply supported, and free edges are presented. The accuracy and validity of the present method is established through convergence and comparison studies with the available literature results.

I. MATHEMATICAL FORMULATION

A. Preliminary

A thick, flat laminated plate with a thickness h, length a, width b, and composed of N orthotropic laminae oriented at angles θ is considered. The reference Cartesian coordinate system is located at the mid-plane of the laminated plate, as depicted in Fig. 1. The laminae are assumed to possess a plane of elastic symmetry parallel to the xy plane and be stacked symmetrically with respect to the middle surface of the laminated plate. The vibration frequencies of the symmetric laminates subjected to a variety of edge conditions, length-to-thickness ratios, aspect ratios, degrees of orthotropy, stacking angles, and numbers of layers are to be determined.

B. Energy expressions

By applying Reddy's higher-order shear deformation theory, the displacements of an arbitrary point of the thick laminated plate along the x, y, and z axes can be represented as

$$u(x,y,z,t)=u_0(x,y,t)+z\phi_x(x,y,t)$$
$$-\frac{4z^3}{3h^2}\left(\phi_x(x,y,t)+\frac{\partial w(x,y,t)}{\partial x}\right), \quad (1a)$$

$$v(x,y,z,t)=v_0(x,y,t)+z\phi_y(x,y,t)$$
$$-\frac{4z^3}{3h^2}\left(\phi_y(x,y,t)+\frac{\partial w(x,y,t)}{\partial y}\right), \quad (1b)$$

$$w(x,y,z,t)=w_0(x,y,t), \quad (1c)$$

where (u_0,v_0,w_0) are the displacement components of the mid-plane along the x, y, and z directions, respectively. The rotations about the x and y directions, respectively, are ϕ_x and ϕ_y.

Assuming transverse inextensibility, the strain–displacement relationship for any lamina in the Cartesian system can be expressed as

$$\varepsilon = L\mathbf{u} \quad (2)$$

with

$$\mathbf{u}=\begin{bmatrix} u_0 & v_0 & w_0 & \phi_x & \phi_y \end{bmatrix}^T, \quad (3a)$$

$$\varepsilon=\begin{bmatrix} \varepsilon_x & \varepsilon_y & \gamma_{yz} & \gamma_{xz} & \gamma_{xy} \end{bmatrix}^T, \quad (3b)$$

and
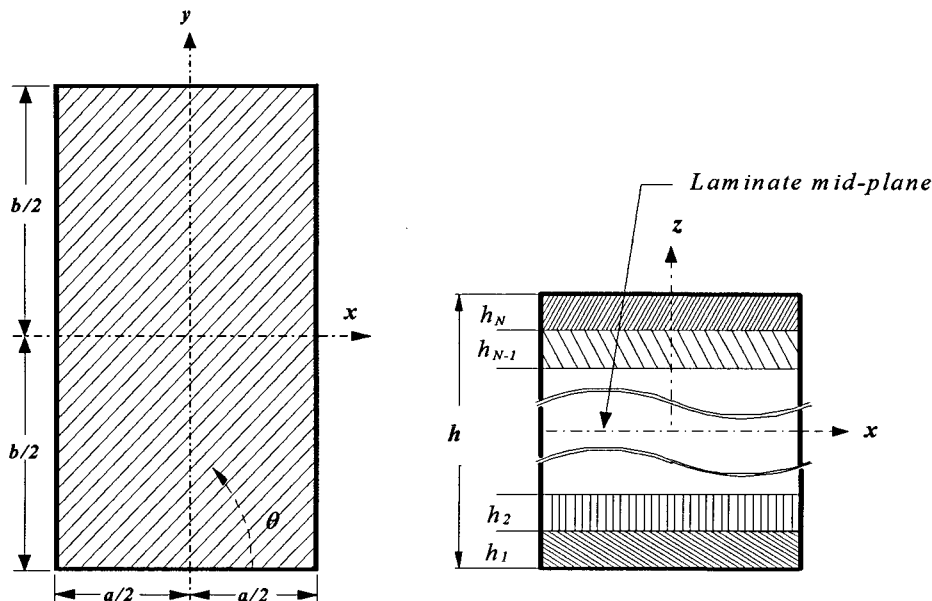
FIG. 1. Geometry of the laminated plate.

$$
L = \begin{bmatrix}
\dfrac{\partial}{\partial x} & 0 & -\dfrac{4z^3 \partial^2}{3h^2\,\partial x^2} & \left(z - \dfrac{4z^3}{3h^2}\right)\dfrac{\partial}{\partial x} & 0 \\[2ex]
0 & \dfrac{\partial}{\partial y} & -\dfrac{4z^3 \partial^2}{3h^2\,\partial y^2} & 0 & \left(z - \dfrac{4z^3}{3h^2}\right)\dfrac{\partial}{\partial y} \\[2ex]
0 & 0 & \left(1 - \dfrac{4z^2}{h^2}\right)\dfrac{\partial}{\partial y} & 0 & \left(1 - \dfrac{4z^2}{h^2}\right) \\[2ex]
0 & 0 & \left(1 - \dfrac{4z^2}{h^2}\right)\dfrac{\partial}{\partial x} & \left(1 - \dfrac{4z^2}{h^2}\right) & 0 \\[2ex]
\dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} & -\dfrac{8z^3 \partial^2}{3h^2\,\partial x\,\partial y} & \left(z - \dfrac{4z^3}{3h^2}\right)\dfrac{\partial}{\partial y} & \left(z - \dfrac{4z^3}{3h^2}\right)\dfrac{\partial}{\partial x}
\end{bmatrix},
\tag{3c}
$$

and the constitutive equations for the kth lamina are

$$
\sigma_k = \mathbf{D}_k \varepsilon_k. \tag{4}
$$

in which $\sigma_k = [\sigma_x\ \sigma_y\ \sigma_{yz}\ \sigma_{xz}\ \sigma_{xy}]^T$ and the stiffness matrix is defined by

$$
\mathbf{D}_k = \begin{bmatrix}
\bar{Q}_{11} & \bar{Q}_{12} & 0 & 0 & \bar{Q}_{16} \\
\bar{Q}_{12} & \bar{Q}_{22} & 0 & 0 & \bar{Q}_{26} \\
0 & 0 & \bar{Q}_{44} & \bar{Q}_{45} & 0 \\
0 & 0 & \bar{Q}_{45} & \bar{Q}_{55} & 0 \\
\bar{Q}_{16} & \bar{Q}_{26} & 0 & 0 & \bar{Q}_{66}
\end{bmatrix}_k. \tag{5}
$$

Here, $(\bar{Q}_{ij})_k$ are obtained by transforming the stacking angle θ and the stiffness constants $(Q_{ij})_k$ which are related to the material properties, E_1, E_2, ν_{12}, ν_{21}, G_{12}, G_{13}, G_{23}, of each ply. In terms of these engineering constants, $(Q_{ij})_k$ can be written in the form of

$$
(Q_{11})_k = E_1/(1 - \nu_{12}\nu_{21}), \tag{6a}
$$

$$
(Q_{12})_k = \nu_{12}E_2/(1 - \nu_{12}\nu_{21}), \tag{6b}
$$

$$
(Q_{22})_k = E_2/(1 - \nu_{12}\nu_{21}), \tag{6c}
$$

$$
(Q_{44})_k = G_{23}, \tag{6d}
$$

$$
(Q_{55})_k = G_{13}, \tag{6e}
$$

$$
(Q_{66})_k = G_{12}. \tag{6f}
$$

In linear elasticity analysis, the strain energy for each ply is given by

$$
U_k = \frac{1}{2} \int_{\mathscr{V}_k} \varepsilon_k^T \sigma_k\, dV_k, \tag{7}
$$

where U_k and V_k are respectively the strain energy and the volume of the kth lamina. Hence, the total strain energy for the entire laminated plate is

$$
U = \frac{1}{2} \sum_{k=1}^{N} \int \int_A \int_{h_{k-1}}^{h_k} (\sigma_x \varepsilon_x + \sigma_y \varepsilon_y + \sigma_{xz}\varepsilon_{xz} + \sigma_{yz}\varepsilon_{yz}
$$
$$
+ \sigma_{xy}\varepsilon_{xy})_k\, dz\, dA. \tag{8}
$$

Accordingly, the total kinetic energy T associated with the vibration of laminated plate is

$$
T = \frac{h}{2} \sum_{k=1}^{N} \int \int_A \rho_k \left[\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial v}{\partial t}\right)^2 + \left(\frac{\partial w}{\partial t}\right)^2 \right] dA \tag{9}
$$

in which ρ_k is the mass density per volume for the kth lamina.

The equivalent modulus for a multidirectional lamina is introduced:

$$
(A_{ij}, B_{ij}, D_{ij}, E_{ij}, F_{ij}, H_{ij})
$$
$$
= \sum_{k=1}^{N} \int_{h_k}^{h_{k+1}} \bar{Q}_{ij}(1, z, z^2, z^3, z^4, z^6)\, dz, \tag{10}
$$

where all B_{ij} and E_{ij} vanish if laminates are stacked symmetrically about the mid-plane. The total potential energy U and total kinetic energy T can be further expanded in terms of the equivalent modulus (see the Appendix). The deflection and rotation functions of the laminate mid-plane are periodic in time. Therefore, for small amplitude vibration, we can assume that

$$
u_0(x,y,t) = U(x,y)\sin \omega t, \tag{11a}
$$

$$
v_0(x,y,t) = V(x,y)\sin \omega t, \tag{11b}
$$

$$
w_0(x,y,t) = W(x,y)\sin \omega t, \tag{11c}
$$

$$
\phi_x(x,y,t) = \Theta_u(x,y)\sin \omega t, \tag{11d}
$$

$$
\phi_y(x,y,t) = \Theta_v(x,y)\sin \omega t. \tag{11e}
$$

By substituting Eqs. (11a)–(11e) into the total potential energy U and the total kinetic energy T, we can obtain the maximum strain energy $U_{\max}$ and the maximum kinetic energy $T_{\max}$. The total energy functional Π of the plate is defined in terms of $U_{\max}$ and $T_{\max}$ as

$$
\Pi = U_{\max} - T_{\max}, \tag{12}
$$

which can be minimized using the p-Ritz method to obtain the vibration frequencies.

C. p-Ritz method

The displacement and rotation components, $U(x,y)$, $V(x,y)$, $W(x,y)$, $\Theta_u(x,y)$, and $\Theta_v(x,y)$, can be further simplified by using the nondimensional expressions

TABLE I. Notations for boundary conditions.

	Boundary constraints[a]	$u_n=u_s=0$	$N_n=u_s=0$	$N_{ns}=u_n=0$	$N_{ns}=N_n=0$
Free	$\partial M_{ns}/\partial s+Q_n=0$ $M_n=0$	F_1	F_2	F_3	F_4
Simply supported	$w=0$ $M_n=0$	S_1	S_2	S_3	S_4
Clamped	$w=0$ $\partial w/\partial n=0$	C_1	C_2	C_3	C_4

[a]Here, n and s indicate the directions normal and tangential to the corresponding supporting edges.

$$\xi=\frac{x}{a}, \tag{13a}$$

$$\eta=\frac{y}{b}. \tag{13b}$$

Accordingly, the in-plane deflection and rotation functions can be expressed in the nondimensional $\xi\eta$ plane, leading to

$$U(\xi,\eta)=\sum_{i=1}^{m} c_i^u \varphi_i^u(\xi,\eta), \tag{14a}$$

$$V(\xi,\eta)=\sum_{i=1}^{m} c_i^v \varphi_i^v(\xi,\eta), \tag{14b}$$

$$W(\xi,\eta)=\sum_{i=1}^{m} c_i^w \varphi_i^w(\xi,\eta), \tag{14c}$$

$$\Theta_u(\xi,\eta)=\sum_{i=1}^{m} c_i^{\theta_u} \varphi_i^{\theta_u}(\xi,\eta), \tag{14d}$$

$$\Theta_v(\xi,\eta)=\sum_{i=1}^{m} c_i^{\theta_v} \varphi_i^{\theta_v}(\xi,\eta), \tag{14e}$$

where φ_i^u, φ_i^v, φ_i^w, $\varphi_i^{\theta_u}$, $\varphi_i^{\theta_v}$ are the so-called p-Ritz functions which are products of two-dimensional polynomials and basic functions. The associated c_i^u, c_i^v, c_i^w, $c_i^{\theta_u}$, $c_i^{\theta_v}$ are the unknown coefficients. The number of terms m in the series (14a)–(14e) can be obtained by

$$m=\frac{(p+1)(p+2)}{2}, \tag{15}$$

where p is the degree of the set of two-dimensional polynomials.

The p-Ritz shape functions can be generated by assuming that

$$\varphi_i^\alpha(\xi,\eta)=\sum_{q=0}^{p}\sum_{j=0}^{q} \xi^{q-j}\eta^j \varphi_b^\alpha(\xi,\eta) \tag{16}$$

in which $\alpha=u,v,w,\theta_u,\theta_v$, and $\varphi_b^\alpha(\xi,\eta)$ denotes the basic function which must at least satisfy the geometric boundary conditions of the laminated plate. The basic function for the laminated plate can be expressed as

$$\varphi_b^\alpha=\prod_{s=1}^{4}[\Gamma_s(\xi,\eta)]^{\Omega_s^\alpha}, \tag{17}$$
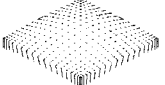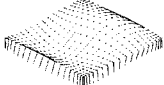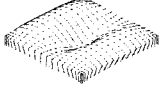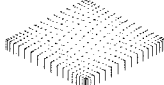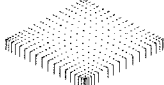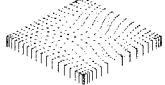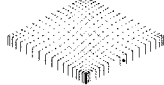
where $\Gamma_s(\xi,\eta)$ is the boundary equation of the sth supporting edge and Ω_s^α denotes the associated basic power. The basic functions consist of products of boundary expressions of the laminated plate raised to their associated basic powers to guarantee automatic satisfaction of geometric boundary conditions. Whitney[13] suggested that one member of each pair of the following four quantities must be prescribed along the boundary to ensure unique solutions to the governing equations

$$u_n;N_n \quad u_s;N_{ns} \quad \frac{\partial w}{\partial n};M_n \quad w;\frac{\partial M_{ns}}{\partial s}+Q_n. \tag{18}$$

TABLE II. Powers of basic functions for various combinations of boundary conditions.

| Edge s | Power Ω_s^a | | | | | | | | | | | | | | | | | | | |
| | 1 | | | | | 2 | | | | | 3 | | | | | 4 | | | | |
α	u	v	w	θ_u	θ_v	u	v	w	θ_u	θ_v	u	v	w	θ_u	θ_v	u	v	w	θ_u	θ_v
$S_2F_4S_2F_4$	0	1	1	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0
$C_1F_4C_1F_4$	1	1	2	1	1	0	0	0	0	0	1	1	2	1	1	0	0	0	0	0
$C_1F_2F_2F_2$	1	1	2	1	1	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0
$C_2F_2F_2F_2$	0	1	2	0	1	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0
$S_2S_2S_2S_2$	0	1	1	0	1	1	0	1	1	0	0	1	1	0	1	1	0	1	1	0
$S_3S_3S_3S_3$	1	0	1	0	1	0	1	1	1	0	1	0	1	0	1	0	1	1	1	0
$C_2S_2S_2S_2$	0	1	2	0	1	1	0	1	1	0	0	1	1	0	1	1	0	1	1	0
$S_2C_1S_2C_1$	0	1	1	0	1	1	1	2	1	1	0	1	1	0	1	1	1	2	1	1
$C_1C_1C_1C_1$	1	1	2	1	1	1	1	2	1	1	1	1	2	1	1	1	1	2	1	1
$C_2C_2C_2C_2$	0	1	2	0	1	1	0	2	1	0	0	1	2	0	1	1	0	2	1	0
$C_3C_3C_3C_3$	1	0	2	1	0	0	1	2	0	1	1	0	2	1	0	0	1	2	0	1

TABLE III. Convergence study of frequency parameters, $\lambda' = \omega h \sqrt{\rho/G}$, for an isotropic thick plate with different degree of polynomial (p).

	Mode shapes						
p							
7	0.0930	0.2220	0.3407	0.4152	0.5256	0.6975	0.7510
9	0.0930	0.2220	0.3406	0.4151	0.5209	0.6846	0.7463
11	0.0930	0.2220	0.3406	0.4151	0.5208	0.6840	0.7454
13	0.0930	0.2220	0.3406	0.4151	0.5208	0.6840	0.7454
15	0.0930	0.2220	0.3406	0.4151	0.5208	0.6840	0.7454
CLPT	0.0963	0.2408	0.3853	0.4816	0.6261	0.8686	0.9632
Mindlin[16]	0.0930	0.2218	0.3402	0.4144	0.5197	0.6821	0.7431
Mallikarjuna[6,a]	0.0929	0.2216	0.3379	0.4184	0.5152	0.6941	0.7610
Noor[17]	0.0932	0.2226	0.3421	0.4171	0.5239	0.6889	0.7511

[a]Higher-order shear deformation theory.

Therefore, many types of boundary conditions are considered along the geometric and natural boundary constraints given in Eq. (18) and summarized in Table I. For examples, there are four types of simply supported conditions: S_1, S_2, S_3, and S_4. These supports have the transverse direction w constrained and varying in-plane support conditions. The most common simply supported condition is S_1 where both the normal and tangential displacements in the midplane are constrained. Another condition having physical interpretation is S_2 with transverse (w) and tangential (u_s) directions constrained which occur mostly in composite plates and sometimes termed "freely supported." [14,15] The basic powers Ω_s^α are assigned 0, 1, or 2 depending on whether the normal, tangential, or transverse direction is constrained at the edge. Details of basic power for various combinations of boundary conditions are listed in Table II. As shown in Fig. 1, $s=1$ refers to the edge at $x=-a/2$, and $s=2,3,4$ correspond to the subsequent edges, going counterclockwise. In detail, the nondimensional basic functions for the rectangular laminated plate are

$$\varphi_b^\alpha = (\xi - 0.5)^{\Omega_1^\alpha}(\eta - 0.5)^{\Omega_2^\alpha}(\xi + 0.5)^{\Omega_3^\alpha}(\eta + 0.5)^{\Omega_4^\alpha} \quad (19)$$

with $\alpha = u, v, w, \theta_u$, and θ_v.

By applying the Ritz method, we minimize the total energy functional Π with respect to the unknown coefficients,

$$\frac{\partial \Pi}{\partial c_i^u} = 0, \tag{20a}$$

$$\frac{\partial \Pi}{\partial c_i^v} = 0, \tag{20b}$$

$$\frac{\partial \Pi}{\partial c_i^w} = 0, \tag{20c}$$

$$\frac{\partial \Pi}{\partial c_i^{\theta_u}} = 0, \tag{20d}$$

$$\frac{\partial \Pi}{\partial c_i^{\theta_v}} = 0, \tag{20e}$$

for i from 1 to m. The differentiation leads to an eigenvalue equation,

$$\{\mathbf{K} - \lambda \mathbf{M}\}\{\mathbf{c}\} = \mathbf{0}, \tag{21}$$

where $\{\mathbf{c}\} = \{c^u \quad c^v \quad c^w \quad ac^{\theta_u} \quad bc^{\theta_v}\}^T$, and λ is the nondimensional eigenvalue defined as

$$\lambda = \frac{\omega ab}{h} \sqrt{\frac{\rho}{D_0}}, \tag{22}$$

with $D_0 = Q_{11}/12$. Details of the stiffness matrices $\mathbf{K}$ and the mass matrices $\mathbf{M}$ are shown in the Appendix.

TABLE IV. Convergence study of frequency parameters, $\lambda' = \omega h \sqrt{\rho/E_2}$, for a three-ply square laminate with different degree of polynomial (p).
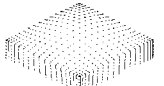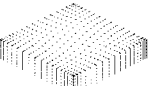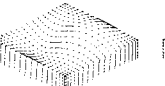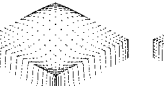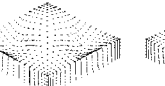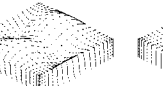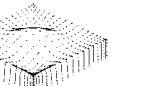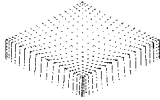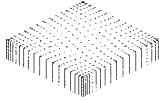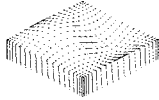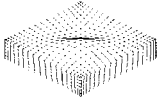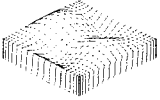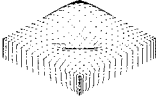
	Mode shapes						
p							
7	0.4537	0.6883	0.7322	0.8449	1.0182	1.1807	1.2653
9	0.4528	0.6883	0.7320	0.8442	1.0173	1.1801	1.2634
11	0.4523	0.6883	0.7320	0.8438	1.0171	1.1799	1.2628
13	0.4520	0.6883	0.7320	0.8435	1.0170	1.1799	1.2624
15	0.4517	0.6883	0.7320	0.8434	1.0170	1.1799	1.2621

TABLE V. Convergence study of frequency parameters, $\lambda' = \omega h \sqrt{\rho/E_2}$, for a five-ply square laminate with different degree of polynomial (p).

	Mode shapes						

p							
7	0.4963	0.6883	0.8085	0.8778	1.1033	1.2468	1.2919
9	0.4955	0.6883	0.8082	0.8772	1.1021	1.2463	1.2907
11	0.4951	0.6883	0.8080	0.8769	1.1017	1.2462	1.2902
13	0.4949	0.6883	0.8080	0.8767	1.1015	1.2462	1.2899
15	0.4948	0.6883	0.8080	0.8765	1.1015	1.2462	1.2897

II. NUMERICAL STUDIES AND DISCUSSIONS

Several examples with various combinations of boundary conditions have been investigated to demonstrate the versatility of the p-Ritz method. All laminae have been assumed to have the same orthotropic properties and equal thicknesses. Numerical results from the published literature have been taken for comparison to demonstrate the accuracy of p-Ritz method integrated with the higher-order shear deformation theory. In this study, all results have been computed in double precision on a *SGI PowerChallenge* computer. Material properties used in all examples have been assumed to be nondimensional as follows:

Material 1: $E_1/E_2 = 1.0$, $G_{12}/E_2 = (1+\nu)/2$,

$$G_{23} = G_{13} = G_{12}, \quad \nu = 0.3.$$

Material 2: $E_1/E_2 = 25$, $G_{12}/E_2 = 0.5$, $G_{23}/E_2 = 0.2$,

$$G_{13} = G_{12}, \quad \nu_{12} = 0.25.$$

Material 3: $E_1/E_2 = 40$, $G_{12}/E_2 = 0.6$, $G_{23}/E_2 = 0.5$,

$$G_{13} = G_{12}, \quad \nu_{12} = 0.25.$$

Material 4: $E_1/E_2 = 1.9\,040\,209$,

$$G_{12}/E_2 = 0.5\,575\,868,$$
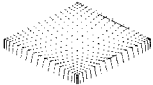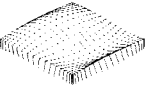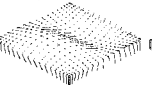
$$G_{23}/E_2 = 0.5\,658\,135,$$

$$G_{13} = 0.3\,391\,225, \quad \nu_{12} = 0.44.$$

The present method has been verified using three examples with an aspect ratio $a/b = 1.0$ and boundary conditions of $S_3S_3S_3S_3$. The first example is an isotropic square plate of material 1 and with a length-to-thickness ratio $a/h = 10$, while the second and the third examples are three-ply and five-ply laminated plates of material 3 with $a/h = 5$ and stacking sequence $(45°/-45°/45°)$ and $(45°/-45°/45°/-45°/45°)$, respectively. A convergence study for frequency parameters was carried out by increasing the number of polynomials from 7 to 15. In Table III, the results for an isotropic plate[16] obtained using this method are very close to Noor's 3-D solutions[17] and Mallikarjuna's solutions[6] which were obtained by a higher-order theory and the finite element method. The frequency parameters obtained using $p = 9$ are within a discrepancy of 0.3% of $p = 15$ in all cases as shown in Tables III–V. It is observed that the fundamental frequencies for single-layer plates converge faster than those for the multi-layer laminates in Tables IV and V. The frequencies also converge rapidly for the modes dominated by in-plane displacements as opposed to the out-of-plane displacements. Higher values for p are therefore needed to provide sufficient accuracy for the higher modes. However, a higher p also leads to more computational effort. Therefore, the parameter $p = 15$ was adopted as a reasonable compromise for subsequent computation.

The fourth example presents a square plate made of ara-

TABLE VI. Comparison of frequency parameters, $\lambda' = \omega h \sqrt{\rho/E_1}$, for an orthotropic square plate made of Material 4.

	Mode shapes						

Source							
CLPT	0.0493	0.1095	0.1327	0.1924	0.2070	0.2671	0.2879	0.3248
Mallikarjuna[6,a]	0.0473	0.1032	0.1190	0.1682	0.1906	0.2205	0.2432	0.2590
Mallikarjuna[6,b]	0.0474	0.1032	0.1190	0.1687	0.1903	0.2201	0.2450	0.2605
Present study	0.0474	0.1032	0.1188	0.1693	0.1885	0.2180	0.2471	0.2623
Reddy[5]	0.0474	0.1033	0.1189	0.1695	0.1888	0.2184	0.2477	0.2629
Srinivas[18]	0.0474	0.1033	0.1188	0.1694	0.1888	0.2180	0.2475	0.2624

[a]First-order shear deformation theory.
[b]Higher-order shear deformation theory.

TABLE VII. Effect of length-to-thickness ratio (a/h) on the fundamental frequency parameters, $\lambda' = (\omega a^2/h)\sqrt{\rho/E_2}$, for four-ply square plates of Material 3.

Source	a/h					
	4	5	10	20	50	100
CLPT	17.907	18.215	18.652	18.767	18.799	18.804
Rikards[7]	9.244	10.690	14.966	17.532	18.632	18.898
Mallikarjuna[6,a]	9.227	10.736	15.073	17.628	18.672	18.835
Mallikarjuna[6,b]	9.258	10.740	15.090	17.637	18.669	18.835
Present study	9.323	10.787	15.107	17.647	18.672	18.836
Reddy[5,b]	9.369	10.820	15.083	17.583	18.590	18.751
Reddy[5,c]	9.497	10.989	15.270	17.668	18.606	18.755

[a]First-order shear deformation theory.
[b]Higher-order shear deformation theory.
[c]Closed form solution.

gonite crystal with $a/h = 10$ and boundary conditions of $S_3S_3S_3S_3$. The mode shapes and frequency parameters are tabulated in Table VI and compared with available results. While the results obtained by the present method have an error of less than 0.25% with respect to both Sriniva's 3-D solutions[18] and Reddy's closed form solutions,[5] the errors in the results from classical laminate plate theory (CLPT), Mallikarjuna[6] (first-order theory), and Mallikarjuna[6] (higher-order theory) are 23.78%, 1.73%, and 1.01% of the exact solutions. As expected, the results obtained by using present method are superior to those of CLPT and first-order theories. They are also very close to the closed form solutions and 3-D elasticity solutions. Therefore, the p-Ritz method is able to provide very accurate results without the difficulties of mesh generations and discretization losses in the finite element method.

In the fifth example, a four-ply square laminate made of material 3 stacked in a sequence of $(0°/90°)_s$ with boundary condition S_2 at all edges and a length-to-thickness ratio of 10 has been examined. The effect of the length-to-thickness ratio on the laminated plates has been presented in Table VII by varying the ratio from 4 to 100. As shown in Table VII, the effect of length-to-thickness ratio on the fundamental frequencies is pronounced and the error in using the classical laminated plate theory increases for thicker plates because of the neglect of shear effects. The present results are in close agreement with Reddy's closed form solutions.

The sixth example analyzed the same laminate as example 5 but with a length-to-thickness ratio of 5, different numbers of layers, and different degree of orthotropy. From the results in Table VIII, it is observed that the prediction of the fundamental frequency by CLPT is inaccurate for materials with a high degree of anisotropy. This reaffirms the fact that the effect of material anisotropy on the fundamental frequency for symmetrically laminated plates is pronounced. In addition, the response characteristics predicted by the present method are accurate and the maximum error in the fundamental frequencies is about 1% compared to the 3-D solutions of Noor.[17]

The effect of boundary conditions on the fundamental frequency has been examined in the seventh example. A four-ply square laminate made of material 3 has been as-

TABLE VIII. Effect of degree of orthotropy on the fundamental frequency, $\lambda' = \omega h \sqrt{\rho/E_2}$, for a square plates with $a/h = 5$ and $S_2S_2S_2S_2$ edge conditions.

Source	Layers	E_1/E_2				
		3	10	20	30	40
CLPT	4	0.2920	0.4126	0.5404	0.6434	0.7320
Reddy[5]		0.2624	0.3309	0.3811	0.4109	0.4315
Present study		0.2624	0.3309	0.3811	0.4109	0.4315
Noor[17]		0.2647	0.3284	0.3824	0.4109	0.4301
CPT	5	0.2920	0.4126	0.5404	0.6434	0.7320
Mallikarjuna[6,a]		0.2626	0.3362	0.3919	0.4246	0.4463
Rikards[7]		0.2608	0.3313	0.3852	0.4142	0.4340
Mallikarjuna[6,b]		0.2626	0.3362	0.3919	0.4248	0.4470
Present study		0.2634	0.3372	0.3937	0.4274	0.4505
Noor[17]		0.2659	0.3409	0.3979	0.4314	0.4537
CPT	9	0.2920	0.4126	0.5404	0.6434	0.7320
Mallikarjuna[6,a]		0.2630	0.3404	0.4011	0.4376	0.4622
Mallikarjuna[6,b]		0.2630	0.3404	0.4011	0.4376	0.4622
Present study		0.2638	0.3413	0.4024	0.4395	0.4648
Noor[17]		0.2664	0.3443	0.4055	0.4421	0.4668

[a]First-order shear deformation theory.
[b]Higher-order shear deformation theory.

TABLE IX. Effect of boundary conditions on the fundamental frequency parameters, $\lambda' = (\omega a^2/h)\sqrt{\rho/E_2}$, for a four-ply square laminate of material 3.

Boundary conditions	Fundamental frequency parameter
$C_2F_2F_2F_2$	1.1378
$F_2F_2C_2F_2$	1.1378
$S_2F_4S_2F_4$	6.2185
$C_1F_4C_1F_4$	12.9422
$S_2S_2S_2S_2$	17.3813
$C_2S_2S_2S_2$	17.6704
$S_2S_2C_2S_2$	17.6704
$S_2C_2S_2C_2$	17.9726
$S_2C_1S_2C_1$	19.5025
$C_1C_1C_1C_1$	21.6722

sumed to have a length-to-thickness ratio of 10 and be stacked with a sequence of $(-45°/45°)_s$. As summarized in Table IX, the frequencies are higher for plates with stiffer constraints.

In the last two examples, the combined effects of plate aspect ratio, length-to-thickness ratio, and stacking angle for symmetric angle-ply, cantilevered ($C_1F_4F_4F_4,$), and simply

TABLE X. Effect of plate aspect ratio (a/b), length-to-thickness ratio (a/h), and lamination angle (θ) on the fundamental frequency parameters, $\lambda' = 100 \times \omega h \sqrt{\rho/E_2}$, for laminated plates of material 3 with stacking sequence as ($\theta/-\theta/\theta/-\theta/\theta$) and $C_1F_4F_4F_4$ edge conditions.

		a/b					
a/h	θ	0.2	0.5	1.0	2.0	5.0	10.0
5	15	15.9392	15.9138	15.7384	15.3293	14.6861	10.6903
	30	13.5254	13.3575	12.6835	11.5457	10.2224	5.8314
	45	10.1767	9.7417	8.5781	7.1679	6.1892	3.1327
	60	6.5221	6.0553	5.3839	4.8222	4.5297	2.2868
	75	4.2761	4.1890	4.1098	4.0619	4.0378	2.0629
	90	3.9612	3.9602	3.9596	3.9593	3.9591	2.0184
100	15	0.0574	0.0572	0.0567	0.0553	0.0525	0.0506
	30	0.0459	0.0451	0.0432	0.0396	0.0338	0.0302
	45	0.0315	0.0302	0.0277	0.0240	0.0192	0.0168
	60	0.0182	0.0172	0.0158	0.0142	0.0125	0.0118
	75	0.0111	0.0109	0.0107	0.0106	0.0104	0.0103
	90	0.0102	0.0102	0.0102	0.0102	0.0102	0.0102

supported ($S_3S_3S_3S_3$) laminated plate have been investigated. The laminates have been assumed to be made of material 3, and be either five-ply with stacking sequence ($\theta/-\theta/\theta/-\theta/\theta$) or three-ply with stacking sequence ($\theta/-\theta/\theta$). The

TABLE XI. Effect of plate aspect ratio (a/b), length-to-thickness ratio (a/h), and lamination angle (θ) on the fundamental frequency parameters, $\lambda' = 100 \times \omega h \sqrt{\rho/E_2}$, for laminated plates with stacking sequence as ($\theta/-\theta/\theta$) and $S_3S_3S_3S_3$ edge conditions.

		a/b							
		Material 2				Material 3			
a/h	θ	0.5	1.0	1.5	2.0	0.5	1.0	1.5	2.0
5	15°	32.2071	35.8155	42.6236	52.0862	37.6402	41.6871	49.7996	61.5597
	30°	29.2007	36.2748	46.6866	58.8802	35.3459	44.0053	56.8605	72.1304
	45°	25.6464	36.6625	50.3993	66.1334	31.6364	45.1748	62.2065	81.8287
	60°	22.1526	36.2748	53.6901	72.9672	26.8135	44.0053	64.5530	87.0976
	75°	18.5448	35.8155	56.3615	78.2015	21.5853	41.6871	64.8946	89.6031
10	15°	11.2081	12.4525	14.9253	18.5448	13.5798	14.8732	17.5509	21.5853
	30°	9.8742	12.6714	16.9729	22.1526	12.1558	15.4325	20.5688	26.8135
	45°	8.3602	12.8639	18.7282	25.6464	10.2682	15.7841	23.0298	31.6364
	60°	6.9411	12.6714	20.3417	29.2007	8.3294	15.4325	24.7607	35.3459
	75°	5.5211	12.4525	21.8693	32.2071	6.4265	14.8732	25.8619	37.6402
20	15°	3.2341	3.6033	4.3657	5.5211	4.0139	4.3921	5.1914	6.4265
	30°	2.8157	3.7035	5.1364	6.9411	3.4941	4.5186	6.2016	8.3294
	45°	2.3346	3.7827	5.8106	8.3602	2.8587	4.6234	7.1143	10.2682
	60°	1.8917	3.7035	6.4131	9.8742	2.2622	4.5186	7.8871	12.1558
	75°	1.4674	3.6033	6.9919	11.2081	1.7097	4.3921	8.5389	13.5798
30	15°	1.4851	1.6566	2.0151	2.5644	1.8549	2.0306	2.4044	2.9870
	30°	1.2901	1.7094	2.3977	3.2813	1.6036	2.0862	2.8919	3.9277
	45°	1.0642	1.7502	2.7378	4.0187	1.3021	2.1367	3.3465	4.9257
	60°	0.8569	1.7094	3.0360	4.8157	1.0240	2.0862	3.7443	5.9624
	75°	0.6607	1.6566	3.3187	5.5135	0.7700	2.0306	4.0998	6.7906
40	15°	0.8455	0.9437	1.1499	1.4674	1.0586	1.1592	1.3739	1.7097
	30°	0.7334	0.9755	1.3748	1.8917	0.9129	1.1907	1.6575	2.2622
	45°	0.6043	0.9999	1.5762	2.3346	0.7392	1.2201	1.9252	2.8587
	60°	0.4854	0.9755	1.7509	2.8157	0.5800	1.1907	2.1616	3.4941
	75°	0.3734	0.9437	1.9152	3.2341	0.4353	1.1592	2.3776	4.0139
50	15°	0.5442	0.6076	0.7411	0.9470	0.6822	0.7472	0.8860	1.1035
	30°	0.4723	0.6287	0.8881	1.2255	0.5876	0.7674	1.0706	1.4649
	45°	0.3886	0.6448	1.0203	1.5185	0.4752	0.7867	1.2459	1.8584
	60°	0.3117	0.6287	1.1343	1.8374	0.3724	0.7674	1.4011	2.2826
	75°	0.2395	0.6076	1.2411	2.1135	0.2792	0.7472	1.5444	2.6337

results are presented in Tables X and XI. In both examples, the fundamental frequencies decrease with an increase in length-to-thickness ratio. It is observed that boundary conditions have a significant influence on the effect of stacking angle for laminated plates. As shown in Table XI, for the simply supported laminates, the fundamental frequencies decrease as stacking angle increases for a plate with $a/b < 1.0$, and increase for $a/b > 1.0$. However, for cantilevered laminates, the fundamental frequencies decrease as the stacking angle increases. It has also been shown that the fundamental frequencies increase with aspect ratio for simply supported laminates and decrease with aspect ratio for cantilevered laminates.

III. CONCLUSIONS

The p-Ritz method has been employed for free vibration analysis of thick composite plates with symmetric lamination based on the higher-order shear deformation theory. A con-cise governing eigenvalue equation has been derived. Convergence of eigenvalues has been verified and excellent agreement has been achieved with respect to first-order, higher-order, finite element, and three-dimensional elasticity solutions. Numerical frequencies for laminates made of materials with different degrees of orthotropy have been presented and illustrated with relevant vibration mode shapes. The effect of length-to-thickness ratio, boundary conditions, plate aspect ratio, number of layers, and stacking angles on the laminates have also been investigated. This analysis suggests that, so far as the free vibration of laminated composite plate is concerned, the higher-order shear deformation theory is able to predict accurate solutions. The versatility of the p-Ritz method in accounting for laminated plates with a variety of boundary constraints should be appreciated.

APPENDIX

Substitution of Eqs. (1a)–(1c), (2), (4), and (10) into Eqs. (8) and (9) yields

$$
\begin{aligned}
U = \frac{1}{2} \int \int_A \Bigg\{ & A_{55}\phi_x^2 + D_{11}\left(\frac{\partial\phi_x}{\partial x}\right)^2 + 2D_{16}\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_x}{\partial y}\right) + D_{66}\left(\frac{\partial\phi_x}{\partial y}\right)^2 + 2A_{45}\phi_x\phi_y + A_{44}\phi_y^2 + 2D_{16}\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_y}{\partial x}\right) \\
& + 2D_{66}\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial\phi_y}{\partial x}\right) + D_{66}\left(\frac{\partial\phi_y}{\partial x}\right)^2 + 2D_{12}\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_y}{\partial y}\right) + 2D_{26}\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial\phi_y}{\partial y}\right) + 2D_{26}\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial\phi_y}{\partial y}\right) + D_{22}\left(\frac{\partial\phi_y}{\partial y}\right)^2 \\
& + F_{44}\left(\frac{16}{h^4}\right)\phi_y^2 + F_{45}\left(\frac{32}{h^4}\right)\phi_x\phi_y + F_{55}\left(\frac{16}{h^4}\right)\phi_x^2 - D_{55}\left(\frac{8}{h^2}\right)\phi_x^2 - D_{45}\left(\frac{16}{h^2}\right)\phi_x\phi_y - F_{16}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_y}{\partial x}\right) - F_{12}\left(\frac{16}{3h^2}\right) \\
& \times \left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_y}{\partial y}\right) - F_{16}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_x}{\partial y}\right) - D_{44}\left(\frac{8}{h^2}\right)\phi_y^2 - F_{11}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial x}\right)^2 - F_{22}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial y}\right)^2 - F_{26}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial y}\right) \\
& \times \left(\frac{\partial\phi_y}{\partial y}\right) - F_{26}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial\phi_y}{\partial y}\right) - F_{66}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial y}\right)^2 + A_{11}\left(\frac{\partial u_0}{\partial x}\right)^2 - F_{66}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial\phi_y}{\partial x}\right) - F_{66}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial x}\right)^2 \\
& + H_{11}\left(\frac{16}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial x}\right)^2 + H_{12}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_y}{\partial y}\right) + H_{16}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial\phi_x}{\partial y}\right) + 2A_{16}\left(\frac{\partial u_0}{\partial x}\right)\left(\frac{\partial v_0}{\partial x}\right) + H_{16}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial x}\right) \\
& \times \left(\frac{\partial\phi_y}{\partial x}\right) + H_{22}\left(\frac{16}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial y}\right)^2 + H_{26}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial\phi_y}{\partial y}\right) + H_{26}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial y}\right)\left(\frac{\partial\phi_y}{\partial x}\right) + H_{66}\left(\frac{16}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial y}\right)^2 \\
& + H_{66}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial\phi_y}{\partial x}\right) + H_{66}\left(\frac{16}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial x}\right)^2 + A_{66}\left(\frac{\partial v_0}{\partial x}\right)^2 + 2A_{16}\left(\frac{\partial u_0}{\partial x}\right)\left(\frac{\partial u_0}{\partial y}\right) + A_{66}\left(\frac{\partial u_0}{\partial y}\right)^2 + 2A_{66}\left(\frac{\partial u_0}{\partial y}\right)\left(\frac{\partial v_0}{\partial x}\right) \\
& + 2A_{12}\left(\frac{\partial u_0}{\partial x}\right)\left(\frac{\partial v_0}{\partial y}\right) + A_{22}\left(\frac{\partial v_0}{\partial y}\right)^2 + 2A_{26}\left(\frac{\partial u_0}{\partial y}\right)\left(\frac{\partial v_0}{\partial y}\right) + 2A_{26}\left(\frac{\partial v_0}{\partial x}\right)\left(\frac{\partial v_0}{\partial y}\right) + 2A_{55}\phi_x\left(\frac{\partial w_0}{\partial x}\right) + 2A_{45}\phi_y\left(\frac{\partial w_0}{\partial x}\right) \\
& + F_{45}\left(\frac{32}{h^4}\right)\phi_y\left(\frac{\partial w_0}{\partial x}\right) + F_{55}\left(\frac{32}{h^4}\right)\phi_x\left(\frac{\partial w_0}{\partial x}\right) - D_{55}\left(\frac{16}{h^2}\right)\phi_x\left(\frac{\partial w_0}{\partial x}\right) - D_{45}\left(\frac{16}{h^2}\right)\phi_y\left(\frac{\partial w_0}{\partial x}\right) + A_{55}\left(\frac{\partial w_0}{\partial x}\right)^2 + F_{55}\left(\frac{16}{h^4}\right)\left(\frac{\partial w_0}{\partial x}\right)^2 \\
& - F_{11}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) + H_{16}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) - F_{12}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) - F_{16}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) \\
& + 2A_{45}\phi_x\left(\frac{\partial w_0}{\partial y}\right) - F_{16}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) + H_{11}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) + 2A_{44}\phi_y\left(\frac{\partial w_0}{\partial y}\right) + H_{12}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) \\
& + H_{16}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right) + H_{11}\left(\frac{16}{9h^4}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right)^2 - D_{44}\left(\frac{8}{h^4}\right)\left(\frac{\partial w_0}{\partial y}\right)^2 - F_{16}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)
\end{aligned}
$$

$$-F_{26}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)-F_{66}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)-F_{66}\left(\frac{16}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)-D_{55}\left(\frac{8}{h^2}\right)\left(\frac{\partial w_0}{\partial x}\right)^2$$

$$+H_{16}\left(\frac{64}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)+H_{26}\left(\frac{64}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)+A_{44}\left(\frac{\partial w_0}{\partial y}\right)^2+H_{66}\left(\frac{64}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)$$

$$+H_{66}\left(\frac{64}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)+F_{44}\left(\frac{16}{h^4}\right)\left(\frac{\partial w_0}{\partial y}\right)^2-H_{16}\left(\frac{64}{9h^4}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)+H_{66}\left(\frac{64}{9h^4}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)^2$$

$$+F_{44}\left(\frac{32}{h^4}\right)\phi_y\left(\frac{\partial w_0}{\partial y}\right)+2A_{45}\left(\frac{\partial w_0}{\partial x}\right)\left(\frac{\partial w_0}{\partial y}\right)+F_{45}\left(\frac{32}{h^4}\right)\left(\frac{\partial w_0}{\partial x}\right)\left(\frac{\partial w_0}{\partial y}\right)-D_{45}\left(\frac{16}{h^2}\right)\left(\frac{\partial w_0}{\partial x}\right)\left(\frac{\partial w_0}{\partial y}\right)-F_{12}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial x}\right)$$

$$\times\left(\frac{\partial^2 w_0}{\partial y^2}\right)-F_{22}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)-D_{45}\left(\frac{16}{h^2}\right)\phi_x\left(\frac{\partial w_0}{\partial y}\right)-F_{26}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_x}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)-F_{26}\left(\frac{8}{3h^2}\right)\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)$$

$$+F_{45}\left(\frac{32}{h^4}\right)\phi_x\left(\frac{\partial w_0}{\partial y}\right)-H_{12}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)+H_{22}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial y}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)-D_{44}\left(\frac{16}{h^2}\right)\phi_y\left(\frac{\partial w_0}{\partial y}\right)+H_{26}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_x}{\partial y}\right)$$

$$\times\left(\frac{\partial^2 w_0}{\partial y^2}\right)+H_{26}\left(\frac{32}{9h^4}\right)\left(\frac{\partial\phi_y}{\partial x}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)+H_{22}\left(\frac{16}{9h^4}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)^2-H_{12}\left(\frac{32}{9h^4}\right)\left(\frac{\partial^2 w_0}{\partial x^2}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)$$

$$+H_{26}\left(\frac{64}{9h^4}\right)\left(\frac{\partial^2 w_0}{\partial x\,\partial y}\right)\left(\frac{\partial^2 w_0}{\partial y^2}\right)\Bigg\}\,dA.\tag{A1}$$

If the laminae are made of the same material with mass density per unit volume ρ, the total kinetic energy becomes

$$T=\frac{\rho\omega^2}{2}\int\int_A\left\{\frac{h^3}{315}\left(17\phi_x^2+17\phi_y^2-8\phi_x\frac{\partial w_0}{\partial x}-8\phi_y\frac{\partial w_0}{\partial y}\right)+h(u_0^2+v_0^2+w_0^2)+\frac{h^3}{252}\left(\phi_x^2+\phi_y^2\right)\right\}dA.\tag{A2}$$

The stiffness matrices $\mathbf{K}$ and the mass matrices $\mathbf{M}$ are given by

$$\mathbf{K}=\frac{1}{D_0}\begin{bmatrix}[\mathbf{K}^{uu}]&[\mathbf{K}^{uv}]&0&0&0\\&[\mathbf{K}^{vv}]&0&0&0\\&&[\mathbf{K}^{ww}]&[\mathbf{K}^{w\theta_u}]&[\mathbf{K}^{w\theta_v}]\\&\text{sym.}&&[\mathbf{K}^{\theta_u\theta_u}]&[\mathbf{K}^{\theta_u\theta_v}]\\&&&&[\mathbf{K}^{\theta_v\theta_v}]\end{bmatrix}\tag{A3}$$

and

$$\mathbf{M}=\begin{bmatrix}[\mathbf{M}^{uu}]&0&0&0&0\\&[\mathbf{M}^{vv}]&0&0&0\\&&[\mathbf{M}^{ww}]&[\mathbf{M}^{w\theta_u}]&[\mathbf{M}^{w\theta_v}]\\&\text{sym.}&&[\mathbf{M}^{\theta_u\theta_u}]&0\\&&&&[\mathbf{M}^{\theta_v\theta_v}]\end{bmatrix},\tag{A4}$$

where the elements of $\mathbf{K}$ and $\mathbf{M}$ can further be expressed as

$$\mathbf{K}_{ij}^{uu}=A_{66}\left(\frac{a^2}{h^3}\right)\mathbf{R}_{\varphi_i^u\varphi_j^u}^{0101}+A_{16}\left(\frac{ab}{h^3}\right)[\mathbf{R}_{\varphi_i^u\varphi_j^u}^{0110}+\mathbf{R}_{\varphi_i^u\varphi_j^u}^{1001}]+A_{11}\left(\frac{b^2}{h^3}\right)\mathbf{R}_{\varphi_i^u\varphi_j^u}^{1010},\tag{A5a}$$

$$\mathbf{K}_{ij}^{uv}=A_{26}\left(\frac{a^2}{h^3}\right)\mathbf{R}_{\varphi_i^u\varphi_j^v}^{0101}+\left(\frac{ab}{h^3}\right)[A_{66}\mathbf{R}_{\varphi_i^u\varphi_j^v}^{0110}+A_{12}\mathbf{R}_{\varphi_i^u\varphi_j^v}^{1001}]+A_{16}\left(\frac{b^2}{h^3}\right)\mathbf{R}_{\varphi_i^u\varphi_j^v}^{1010},\tag{A5b}$$

$$\mathbf{K}_{ij}^{vv}=A_{22}\left(\frac{a^2}{h^3}\right)\mathbf{R}_{\varphi_i^v\varphi_j^v}^{0101}+A_{26}\left(\frac{ab}{h^3}\right)[\mathbf{R}_{\varphi_i^v\varphi_j^v}^{0110}+\mathbf{R}_{\varphi_i^v\varphi_j^v}^{1001}]+A_{66}\left(\frac{b^2}{h^3}\right)\mathbf{R}_{\varphi_i^v\varphi_j^v}^{1010},\tag{A5c}$$

$$
\mathbf{K}_{ij}^{ww} = \left[A_{44}\left(\frac{a^2}{h^3}\right) - D_{44}\left(\frac{8a^2}{h^5}\right) + F_{44}\left(\frac{16a^2}{h^7}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^w}^{0101} + \left[A_{45}\left(\frac{ab}{h^3}\right) - D_{45}\left(\frac{8ab}{h^5}\right) + F_{45}\left(\frac{16ab}{h^7}\right) \right] \left[\mathbf{R}_{\varphi_i^w \varphi_j^w}^{0110} + \mathbf{R}_{\varphi_i^w \varphi_j^w}^{1001} \right] + \left[A_{55}\left(\frac{b^2}{h^3}\right) \right.
$$
$$
\left. - D_{55}\left(\frac{8b^2}{h^5}\right) + F_{55}\left(\frac{16b^2}{h^7}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^w}^{1010} + H_{22}\left(\frac{16a^2}{9b^2h^7}\right) \mathbf{R}_{\varphi_i^w \varphi_j^w}^{0202} + H_{26}\left(\frac{32a}{9bh^7}\right) \left[\mathbf{R}_{\varphi_i^w \varphi_j^w}^{0211} + \mathbf{R}_{\varphi_i^w \varphi_j^w}^{1102} \right] + H_{66}\left(\frac{64}{9h^7}\right) \mathbf{R}_{\varphi_i^w \varphi_j^w}^{1111}
$$
$$
+ H_{16}\left(\frac{32b}{9ah^7}\right) \left[\mathbf{R}_{\varphi_i^w \varphi_j^w}^{1120} + \mathbf{R}_{\varphi_i^w \varphi_j^w}^{2011} \right] + H_{12}\left(\frac{16}{9h^7}\right) \left[\mathbf{R}_{\varphi_i^w \varphi_j^w}^{0220} + \mathbf{R}_{\varphi_i^w \varphi_j^w}^{2002} \right] + H_{11}\left(\frac{16b^2}{9a^2h^7}\right) \mathbf{R}_{\varphi_i^w \varphi_j^w}^{2020}, \tag{A5d}
$$

$$
\mathbf{K}_{ij}^{w\theta_u} = \left[A_{45}\left(\frac{ab}{h^3}\right) - D_{45}\left(\frac{8ab}{h^5}\right) + F_{45}\left(\frac{16ab}{h^7}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{0100} + \left[H_{26}\left(\frac{16a}{9bh^7}\right) - F_{26}\left(\frac{4a}{3bh^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{0201} + \left[H_{12}\left(\frac{16}{9h^7}\right) \right.
$$
$$
\left. - F_{12}\left(\frac{4}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{0210} + \left[A_{55}\left(\frac{b^2}{h^3}\right) - D_{55}\left(\frac{8b^2}{h^5}\right) + F_{55}\left(\frac{16b^2}{h^7}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{1000} + \left[H_{66}\left(\frac{32}{9h^7}\right) - F_{66}\left(\frac{8}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{1101}
$$
$$
+ \left[H_{16}\left(\frac{16b}{9ah^7}\right) - F_{16}\left(\frac{4b}{3ah^5}\right) \right] \left[2\mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{1110} + \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{2001} \right] + \left[H_{11}\left(\frac{16b^2}{9a^2h^7}\right) - F_{11}\left(\frac{4b^2}{3a^2h^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{2010}, \tag{A5e}
$$

$$
\mathbf{K}_{ij}^{w\theta_v} = \left[A_{44}\left(\frac{a^2}{h^3}\right) - D_{44}\left(\frac{8a^2}{h^5}\right) + F_{44}\left(\frac{16a^2}{h^7}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{0100} + \left[H_{22}\left(\frac{16a^2}{9b^2h^7}\right) - F_{22}\left(\frac{4a^2}{3b^2h^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{0201} + \left[H_{26}\left(\frac{16a}{9bh^7}\right) \right.
$$
$$
\left. - F_{26}\left(\frac{4a}{3bh^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{0210} + \left[A_{45}\left(\frac{ab}{h^3}\right) - D_{45}\left(\frac{8ab}{h^5}\right) + F_{45}\left(\frac{16ab}{h^7}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{1000} + \left[H_{26}\left(\frac{32a}{9bh^7}\right) - F_{26}\left(\frac{8a}{3bh^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{1101}
$$
$$
+ \left[H_{66}\left(\frac{32}{9h^7}\right) - F_{66}\left(\frac{8}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{1110} + \left[H_{12}\left(\frac{16}{9h^7}\right) - F_{12}\left(\frac{4}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{2001} + \left[H_{16}\left(\frac{16b}{9ah^7}\right) - F_{16}\left(\frac{4b}{3ah^5}\right) \right] \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{2010}, \tag{A5f}
$$

$$
\mathbf{K}_{ij}^{\theta_u \theta_u} = \left[A_{55}\left(\frac{b^2}{h^3}\right) - D_{55}\left(\frac{8b^2}{h^5}\right) + F_{55}\left(\frac{16b^2}{h^7}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_u}}^{0000} + \left[D_{66}\left(\frac{1}{h^3}\right) + H_{66}\left(\frac{16}{9h^7}\right) - F_{66}\left(\frac{8}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_u}}^{0101} + \left[D_{16}\left(\frac{b}{ah^3}\right) \right.
$$
$$
\left. + H_{16}\left(\frac{16b}{9ah^7}\right) - F_{16}\left(\frac{8b}{3ah^5}\right) \right] \left[\mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_u}}^{0110} + \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_u}}^{1001} \right] + \left[D_{11}\left(\frac{b^2}{a^2h^3}\right) + H_{11}\left(\frac{16b^2}{9a^2h^7}\right) - F_{11}\left(\frac{8b^2}{3a^2h^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_u}}^{1010}, \tag{A5g}
$$

$$
\mathbf{K}_{ij}^{\theta_u \theta_v} = \left[A_{45}\left(\frac{ab}{h^3}\right) - D_{45}\left(\frac{8ab}{h^5}\right) + F_{45}\left(\frac{16ab}{h^7}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_v}}^{0000} + \left[D_{26}\left(\frac{a}{bh^3}\right) + H_{26}\left(\frac{16a}{9bh^7}\right) - F_{26}\left(\frac{8a}{3bh^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_v}}^{0101}
$$
$$
+ \left[D_{66}\left(\frac{1}{h^3}\right) + H_{66}\left(\frac{16}{9h^7}\right) - F_{66}\left(\frac{8}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_v}}^{0110} + \left[D_{12}\left(\frac{1}{h^3}\right) + H_{12}\left(\frac{16}{9h^7}\right) - F_{12}\left(\frac{8}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_v}}^{1001}
$$
$$
+ \left[D_{16}\left(\frac{b}{ah^3}\right) + H_{16}\left(\frac{16b}{9ah^7}\right) - F_{16}\left(\frac{8b}{3ah^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_v}}^{1010}, \tag{A5h}
$$

$$
\mathbf{K}_{ij}^{\theta_v \theta_v} = \left[A_{44}\left(\frac{a^2}{h^3}\right) - D_{44}\left(\frac{8a^2}{h^5}\right) + F_{44}\left(\frac{16a^2}{h^7}\right) \right] \mathbf{R}_{\varphi_i^{\theta_v} \varphi_j^{\theta_v}}^{0000} + \left[D_{22}\left(\frac{a^2}{b^2h^3}\right) + H_{22}\left(\frac{16a^2}{9b^2h^7}\right) - F_{22}\left(\frac{8a^2}{3b^2h^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_v} \varphi_j^{\theta_v}}^{0101} + \left[D_{26}\left(\frac{a}{bh^3}\right) \right.
$$
$$
\left. + H_{26}\left(\frac{16a}{9bh^7}\right) - F_{26}\left(\frac{8a}{3bh^5}\right) \right] \left[\mathbf{R}_{\varphi_i^{\theta_v} \varphi_j^{\theta_v}}^{0110} + \mathbf{R}_{\varphi_i^{\theta_v} \varphi_j^{\theta_v}}^{1001} \right] + \left[D_{66}\left(\frac{1}{h^3}\right) + H_{66}\left(\frac{16}{9h^7}\right) - F_{66}\left(\frac{8}{3h^5}\right) \right] \mathbf{R}_{\varphi_i^{\theta_v} \varphi_j^{\theta_v}}^{1010}, \tag{A5i}
$$

and

$$
\mathbf{M}_{ij}^{uu} = h\mathbf{R}_{\varphi_i^u \varphi_j^u}^{0000}, \tag{A6a}
$$

$$
\mathbf{M}_{ij}^{vv} = h\mathbf{R}_{\varphi_i^v \varphi_j^v}^{0000}, \tag{A6b}
$$

$$
\mathbf{M}_{ij}^{ww} = h\mathbf{R}_{\varphi_i^w \varphi_j^w}^{0000} + \left(\frac{h^3}{252b^2}\right) \mathbf{R}_{\varphi_i^w \varphi_j^w}^{0101} + \left(\frac{h^3}{252a^2}\right) \mathbf{R}_{\varphi_i^w \varphi_j^w}^{1010}, \tag{A6c}
$$

$$
\mathbf{M}_{ij}^{w\theta_u} = \left(\frac{-4h^3}{315a^2}\right) \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_u}}^{1000}, \tag{A6d}
$$

$$
\mathbf{M}_{ij}^{w\theta_v} = \left(\frac{-4h^3}{315b^2}\right) \mathbf{R}_{\varphi_i^w \varphi_j^{\theta_v}}^{0100}, \tag{A6e}
$$

$$
\mathbf{M}_{ij}^{\theta_u \theta_u} = \left(\frac{17h^3}{315a^2}\right) \mathbf{R}_{\varphi_i^{\theta_u} \varphi_j^{\theta_u}}^{0000}, \tag{A6f}
$$

$$
\mathbf{M}_{ij}^{\theta_v \theta_v} = \left(\frac{17h^3}{315b^2}\right) \mathbf{R}_{\varphi_i^{\theta_v} \varphi_j^{\theta_v}}^{0000}, \tag{A6g}
$$

in which

$$\mathbf{R}^{defg}_{\varphi_i^\alpha \theta_j^\beta} = \int \int_A \frac{\partial^{d+e} \varphi_i^\alpha(\xi,\eta)}{\partial \xi^d \, \partial \eta^e} \frac{\partial^{f+g} \theta_j^\beta(\xi,\eta)}{\partial \xi^f \, \partial \eta^g} \, d\xi \, d\eta. \tag{A7}$$

[1] P. C. Yang, C. H. Norris, and Y. Stavsky, ''Elastic wave propagation in heterogeneous plates,'' Int. J. Solids Struct. **2**, 665–684 (1966).

[2] J. N. Reddy, ''A simple higher-order theory for laminated composite plates,'' J. Appl. Mech. ASME **51**, 745–752 (1984).

[3] J. N. Reddy and D. H. Robbins, Jr., ''Theories and computational models for laminated composite laminates,'' Appl. Mech. Rev. **47**(6), 147–169 (1994).

[4] A. K. Noor and W. S. Burton, ''Assessment of shear deformation theories for multilayered composite plates,'' Appl. Mech. Rev. **42**(1), 1–13 (1989).

[5] J. N. Reddy and N. D. Phan, ''Stability and vibration of isotropic, orthotropic and laminated plates according to a higher-order shear deformation theory,'' J. Sound Vib. **98**(2), 157–170 (1985).

[6] Mallikarjuna and T. Kant, ''Free vibration of symmetrically laminated plates using a higher-order theory with finite element technique,'' Int. J. Numer. Methods Eng. **28**, 1875–1889 (1989).

[7] R. Rikards, A. Chate, and A. Korjakin, ''Vibration and damping analysis of laminated composite plates by the finite element method,'' Eng. Comput. **12**(1), 61–74 (1995).

[8] K. M. Liew and C. M. Wang, ''pb-2 Rayleigh-Ritz method for general plate analysis,'' Eng. Struct. **15**(1), 55–60 (1993).

[9] C. W. Lim and K. M. Liew, ''Effects of boundary constraints and thickness variations on the vibratory response of rectangular plates,'' Thin-Walled Struct. **17**(2), 133–159 (1993).

[10] K. M. Liew and C. W. Lim, ''Vibratory characteristics of general laminates, I: symmetric trapezoids,'' J. Sound Vib. **183**(4), 615–642 (1995).

[11] K. M. Liew, Y. Xiang, and S. Kitipornchai, ''Research on thick plate vibration: a literature survey,'' J. Sound Vib. **180**(1), 163–176 (1995).

[12] K. M. Liew, ''Solving the vibration of thick symmetric laminates by Reissner/Midlin plate theory and the *p*-Ritz method,'' J. Sound Vib. **198**(3), 343–360 (1996).

[13] J. M. Whitney, *Structural Analysis of Laminated Anisotropic Plates* (Technomic, Lancaster, PA, 1987), pp. 39–40.

[14] B. Baharlou and A. W. Leissa, ''Vibration and buckling of generally laminated composite plates with arbitrary edge conditions,'' Int. J. Mech. Sci. **29**(8), 545–555 (1987).

[15] A. W. Leissa, *Vibration of shells* (NASA SP-288) (U.S. Government Printing Office, Washington DC, 1973), p. 43.

[16] R. D. Mindlin, ''Influence of rotary inertia and shear on flexural motions of isotropic, elastic plates,'' J. Appl. Mech. **18**, 31–38 (1951).

[17] A. K. Noor, ''Free vibrations of multilayered composite plates,'' AIAA J. **11**, 1038–1039 (1973).

[18] S. Srinivas and A. K. Rao, ''Bending, vibration and buckling of simply supported thick orthotropic rectangular plates and laminates,'' Int. J. Solids Struct. **6**(8), 1463–1481 (1970).

Sensitivity of structural acoustic response to attachment feature scales

W. Steve Shepard, Jr. and Kenneth A. Cunefare
School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

When implementing a discrete computational method to predict the vibro-acoustic behavior of a larger structure one must decide how much effort and refinement to expend on modeling smaller attached structural components. In representing these attachments, or features, it is important to know what scales are important with regard to a valid prediction of the system response. To aid in the study of the impact of feature scales, sensitivity relationships for important system variables are developed in this paper. The sensitivity analysis will, for one, consider the change in the radiated power with respect to changes in the scale of an attached mass. These relations are developed analytically and do not require finite difference methods or eigenvalue derivatives. It is shown that computing the sensitivity of the radiated acoustic power only requires evaluating one additional term as all of the other terms are found during the normal course of modeling the system. Furthermore, the terms that include fluid-loading on the structure need only be reevaluated when the excitation frequency is changed. It is also noted that the sensitivity relations developed here can easily be used to investigate the effects of modifying a number of structural parameters. © *1997 Acoustical Society of America.* [S0001-4966(97)02005-5]

INTRODUCTION

Most structural systems feature a variety of structural subsystems or subcomponents. When predicting the dynamic and acoustic behavior of a system one must decide how much effort will be spent on modeling these smaller structural features. Consider, for example, a structural system that contains a simple mass element. The question arises, when does the size of this mass become large enough to justify a greater modeling effort? That is, can the mass be modeled using a lumped representation or must it be treated using some distributed representation? Moreover, how accurate must a spatial representation be to obtain a valid acoustic prediction? Researchers often encounter conditions when it is necessary to improve the modeling of an attached structural feature. For example, Ettouney *et al.*[1] observed that it was important to model the actual geometry of an attached mass to obtain an accurate solution for the shell structure analyzed in their research. A more recent work that examines the impact of various distribution functions on the scattering of sound from an infinite fluid-loaded plate is that of Feit and Cuschieri.[2] In their paper, several distribution functions were used to represent a mass inhomogeneity. The effect of using each of these distributions on the predicted surface velocity and far-field scattered pressure was examined. However, the width of the attachment remained unchanged throughout the analysis. These works have focused on how the presence of a feature modifies the system behavior. The affect of the feature scale on the problem was not addressed. A comprehensive study of scales that are critical to feature modeling in structural-acoustic systems has not been addressed.

The purpose of this work is to study the impact of feature scales. To that end, relationships are developed that express the sensitivity of important system variables to changes in a feature's scale. The variables examined are the surface displacement, surface velocity, surface pressure, and the radiated acoustic power. The feature considered in this paper is an attached mass. Other researchers have developed sensitivity analyses for various system parameters. The sensitivity of system eigenvalues with respect to mass location was considered by Wang.[3] St. Pierre and Koopmann determined the sensitivity of the radiated power to changes in the amount of mass attached at a point on a rectangular plate.[4] However, changes in the shape or size of the mass were not considered in either of these works. Varying the material properties of a structure to minimize the radiated acoustic power has also been studied.[5] Furthermore, the sensitivity of the radiated power to various design or modeling parameters has more recently been presented by Salagame *et al.*[6] Unlike the work of Salagame *et al.*, however, the effects of fluid-loading on the structure are included in the present work. Furthermore, the impedance matrix used in the present work depends on the parameter being changed. The impact of an arbitrary impedance matrix perturbation on the coupled vibro-acoustic response has been investigated by Hahn and Ferri.[7,8] In their work, the inverse of the system matrix was approximated by neglecting the higher-order terms. Due to the particular approach taken here, however, the inverse of this matrix will be computed without approximations. It will also be seen that all of the terms needed to compute these sensitivities, except one, will have been determined in the normal course of modeling the system.

Further insight into the physics of the problem will be gained by decomposing the surface response into ''radiation modes.'' The advantage of using these modes is that their shapes do not depend on the internal structure of the system. A discussion of properties of radiation modes, along with a review of research in the area, is given by Cunefare and
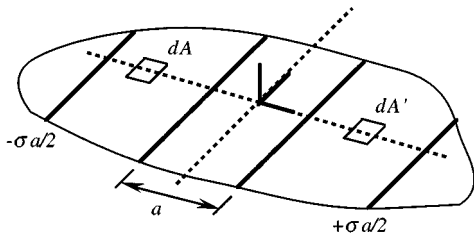
FIG. 1. Semi-infinite fluid-loaded elastic plate in rigid baffle.

Currey.[9] The results presented in a recent paper for fluid-loaded structures[10] are applied here.

Once the influence of feature scales is identified, a greater understanding of how the size and modeling detail of a particular feature can affect the acoustic output and vibratory response of a system will be gained. Furthermore, future modeling efforts of structural-acoustic systems can be directed in a more efficient manner.

The following development is for a harmonically excited fluid-loaded plate that carries an attached mass. Since closed form solutions to this type of problem are not readily available, the acoustic surface variational principle (SVP)[11] is used to model the system. The SVP is ideally suited to this research since it does not implement a surface discretization that has its own set of spatial scales. Note, however, that the numerical techniques applied in this research do not restrict the application of the results to other methods. The following section reviews the system relations for the bare fluid-loaded plate. The modeling of the attached mass feature is then presented. Finally, sensitivity relations for important system variables are given along with a numerical example.

I. CLEAN SYSTEM MODELING

This section provides a brief review of the system equations for a semi-infinite plate that is simply supported in a rigid baffle and has one wetted surface as shown in Fig. 1. For details in deriving the system of equations for this system, refer to Ref. 12.

The surface pressure and surface velocity are nondimensionalized, respectively, as

$$\hat{p}(\xi)=\frac{1}{\rho c^2}p(\xi), \quad \hat{\nu}_n(\xi)=\frac{1}{c}\nu_n(\xi), \quad \xi=\frac{x}{a}, \quad (1)$$

and ξ is the nondimensional position. The density ρ and sound speed c are for the acoustic fluid and a is the width of the plate. The carat will now be dropped since all the terms are nondimensional. The mass per unit surface area of the plate is $\rho_s h$, where h is the plate thickness and ρ_s is the plate density. Furthermore, the characteristic wave speed for the structure, $c_s=D/\rho_s h a^2$, depends on the bending stiffness of the plate, $D=Eh^3/[12(1-\nu^2)]$, where E is the elastic modulus and ν is Poisson's ratio.

The surface-pressure distribution is expressed over a region of the baffle, $-\sigma/2\leq\xi\leq\sigma/2$, using a Ritz expansion with N linearly independent basis functions as

$$p(\xi)=\sum_{k=1}^{N}\psi_k(\xi)P_k \quad \text{where} \quad \psi_k(\xi)=0, \quad \xi=\frac{\sigma}{2} \quad (2)$$

and it is assumed the pressure vanishes beyond the distance $r=\sigma a/2$.[13,14] That is, the surface pressure is zero for $\xi<-\sigma/2$ and $\xi>\sigma/2$. It is only necessary that the functions, ψ, be both continuous and piecewise differentiable. Similarly, the surface displacement is expressed over the plate surface $-1/2\leq\xi\leq1/2$ as

$$w(\xi,t)=a\sum_{j=1}^{M}\phi_j(\xi)q_j(t) \quad \text{where} \quad \phi_j(\xi)=0, \quad \xi=\pm\frac{1}{2} \quad (3)$$

such that the basis functions, ϕ, are continuous and piecewise differentiable. It is important to make N larger than M by a factor of σ (i.e., $N=\sigma M$). A harmonic time dependence is assumed for the generalized modal coordinate

$$q_j(t)=\text{Re}[W_j\exp(-i\omega t)]. \quad (4)$$

Therefore, the nondimensional normal surface velocity is

$$\nu_n(\xi,t)=-ika\sum_{j=1}^{M}\phi_j(\xi)q_j(t), \quad (5)$$

where ω is the frequency of excitation and $ka=a\omega/c$ is the nondimensional wave number. By applying the surface variational principle, a system of equations is obtained relating the unknown modal pressure amplitudes P_k and unknown displacement amplitudes W_j

$$[A]\{P\}=2\pi(ka)^2[R]^T\{W\}, \quad (6)$$

where

$$A_{jl}=\int_{-\sigma/2}^{\sigma/2}\int_{-\sigma/2}^{\sigma/2}\{(ka)\psi_j(\xi)\psi_l(\zeta)$$
$$-\psi_j'(\xi)\psi_l'(\zeta)\}G(\zeta|\xi)d\zeta\,d\xi \quad (7)$$

and

$$R_{jl}=\int_{-\sigma/2}^{\sigma/2}\phi_j\psi_l\,d\xi. \quad (8)$$

The superscript T refers to the matrix transpose and in Eq. (7), $G(\zeta|\xi)$ is the two-dimensional Green's function. Note that while R_{jl} is purely real, A_{jl} is complex.

The remaining equations are the coupled equations for structural vibration. These relations are found by applying Lagrange's equation to the potential and kinetic energy expressions for the plate.[12] The resulting system of equations is

$$[R]\{P\}+[D]\{W\}=\{F^e\}, \quad (9)$$

where

$$D_{jk}=\frac{\rho_s h}{\rho a}\left[\left(\frac{c_s}{c}\right)^2\kappa_{jk}-(ka)^2\mu_{jk}\right] \quad (10)$$

and the expression for R_{jk} is given in Eq. (8). The nondimensional generalized inertia and stiffness coefficients are, respectively,

$$\mu_{jl}=\int_{-1/2}^{1/2}\phi_j\phi_l\,d\xi \quad (11)$$

and
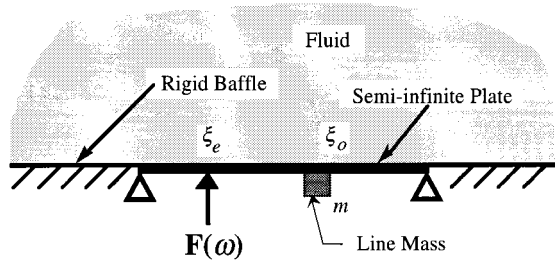
FIG. 2. Inertial Scales: Semi-infinite plate with line mass attachment.

$$\kappa_{jl} = \int_{-1/2}^{1/2} \phi_j'' \phi_l'' \, d\xi. \tag{12}$$

A harmonic forcing function is assumed to excite the system. As such, the components of the excitation vector $\{F^e\}$ are found from

$$F_j = \int_{-1/2}^{1/2} f(\xi) \phi_j \, d\xi. \tag{13}$$

For the work presented here, the harmonic forcing function is a line excitation, $f(\xi) = f\delta(\xi - \xi_e)$, where ξ_e is the location of the force of magnitude f.

Combining the Lagrangian equations, Eq. (9), and the acoustic equations, Eq. (6), a square system of equations is obtained which completely describes the coupled fluid-structure response. Once these relations are solved for the modal coefficients, P_k and W_j, the surface pressure and displacement can be constructed from Eqs. (2) and (3), respectively.

II. ATTACHED FEATURE MODELING

The effects of inertial scales are investigated by considering a line mass attached to the semi-infinite plate as shown in Fig. 2. For the current study, the mass is modeled using a chapeau representation. Denoting the unit step at ξ by $u(\xi)$, the mass is represented by

$$m_c(\xi) = \frac{m_{rat}}{\Delta_m} [u(\xi - (\xi_0 - \Delta_m/2)) - u(\xi - (\xi_0 + \Delta_m/2))], \tag{14}$$

where m_{rat} is the total added mass divided by the plate mass, $m_{rat} = m_{added}/\rho_s h a$. The scale of the attached mass can be changed by appropriately changing the width of the distribution Δ_m. A few example distributions are shown in Fig. 3 for a mass located at $\xi = 0.2071$. As Δ_m decreases the mass becomes more centralized around the attachment position, ξ_0. Conversely, as Δ_m increases the mass becomes more distributed over the plate. Note that since it is desired to only compare scales, when the value of Δ_m is changed the total added mass remains constant.

For comparison purposes, the mass will also be modeled using a lumped representation that confines all of the mass to a single line parallel with the edge of the plate

$$m_p(\xi) = m_{rat}\delta(\xi - \xi_0), \tag{15}$$

where $\delta(\xi - \xi_0)$ represents the Dirac delta function at ξ_0.
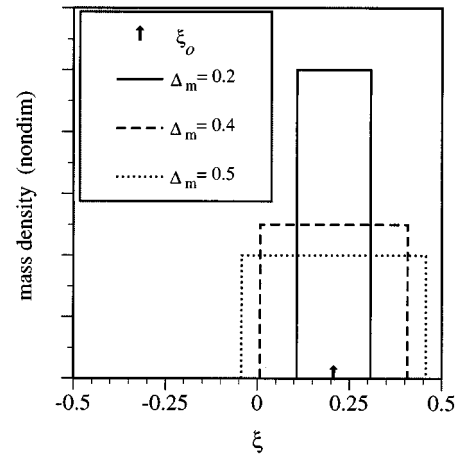
FIG. 3. Sample chapeau distributions for mass located at $\xi_0 = 0.2071$.

Regardless of the distribution used, the mass is incorporated into the system formulation by determining the nondimensional kinetic energy due to the added mass alone,

$$T_m = \frac{1}{2} \frac{\rho_s h}{\rho a} \frac{a^2}{c^2} \sum_{j=1}^{M} \sum_{k=1}^{M} \int_{-1/2}^{1/2} m(\xi) \phi_j \phi_k \, d\xi \, \dot{q}_j q_k. \tag{16}$$

Recall that $m(\xi)$ is relative to the plate mass in Eqs. (14) and (15). After applying Lagrange's equations, the relation for D_{jk} in Eq. (10) is modified as

$$D_{jk}^m = \frac{\rho_s h}{\rho a} \left[\left(\frac{c_s}{c} \right)^2 \kappa_{jk} - (ka)^2 \mu_{jk} - (ka)^2 \eta_{jk} \right], \tag{17}$$

where the inertial coupling coefficient, η_{jk}, depends on the type of distribution used. For the chapeau distribution in Eq. (14), the inertial coupling coefficient is

$$\eta_{jk} = \frac{m_{rat}}{\Delta_m} \int_{\xi_0 - \Delta_m/2}^{\xi_0 + \Delta_m/2} \phi_j \phi_k \, d\xi. \tag{18}$$

Whereas, when the lumped representation, Eq. (15), is used,

$$\eta_{jk} = m_{rat}\phi_j(\xi_0)\phi_k(\xi_0). \tag{19}$$

The forced response of the combined system can now be predicted using Eqs. (6) and (9) provided D_{jk} is replaced by D_{jk}^m. The resulting $M+N$ by $M+N$ system of equations is

$$\begin{bmatrix} [A] & -2\pi(ka)^2[R]^T \\ [R] & [D^m] \end{bmatrix} \begin{Bmatrix} P \\ W \end{Bmatrix} = \begin{Bmatrix} \{0\} \\ \{F^e\} \end{Bmatrix}. \tag{20}$$

It should be noted that in the formulation presented here the attached mass is distributed as a locally reacting structure in that it can only exert forces normal to the surface of the plate. In other words, the mass does not react to bending forces.

III. SENSITIVITY OF SYSTEM VARIABLES TO FEATURE SCALES

This research focuses on the influence of inertial scales on the structural-acoustic response of a semi-infinite plate. Scales associated with the structural element, or feature, were introduced into the numerical formulation by describing the element spatial distribution with a chapeau function.

It is now desired to determine the sensitivities of important system variables to changes in this scale. Although the method developed here is specifically related to changes in the inertial scale, the results can easily be applied to changes in any structural parameter. Such extensions are briefly noted at the end of this section.

Several metrics could be used to determine the effects of the inertial scale. However, those considered here are the plate surface displacement; surface velocity, surface pressure, and radiated acoustic power. Relations that express the sensitivity of these variables to changes in the inertial scale are developed. To facilitate such a development, the system of equations developed in the previous section is first rearranged into a more convenient form.

By solving the first line in Eq. (20) for $\{P\}$ and substituting the result into the second line, a relation containing an impedance matrix is obtained

$$[Z]_{\mathrm{sys}}\{W\}=\{F^e\}. \tag{21}$$

The combined system impedance matrix $[Z]_{\mathrm{sys}}$ of size $M \times M$ is given by

$$[Z]_{\mathrm{sys}}=[Z]_{\mathrm{fl}}+[Z]_{\mathrm{st}}, \tag{22}$$

where the fluid-loading impedance matrix is

$$[Z]_{\mathrm{fl}}=2\pi(ka)^2[R][A]^{-1}[R]^T, \tag{23}$$

and the structural impedance matrix is

$$[Z]_{\mathrm{st}}=[D^m]. \tag{24}$$

It is important at this point to note that only the structural impedance matrix, $[Z]_{\mathrm{st}}$, depends on the scale of the feature.

First, consider the sensitivity of the surface displacement coefficients to changes in the mass width. Solving Eq. (21) for $\{W\}$ and taking the derivative with respect to Δ_m, we obtain

$$\frac{\partial\{W\}}{\partial\Delta_m}=\frac{\partial}{\partial\Delta_m}\left([Z]_{\mathrm{sys}}^{-1}\right)\{F^e\}+[Z]_{\mathrm{sys}}^{-1}\frac{\partial\{F^e\}}{\partial\Delta_m}. \tag{25}$$

However, since the forcing vector is not a function of Δ_m, the second term in Eq. (25) is zero. Furthermore, since $[Z]_{\mathrm{sys}}^{-1}[Z]_{\mathrm{sys}}=[I]$ and $\partial[I]/\partial\Delta_m=0$, the identity (see Ref. 15, p. 43)

$$\frac{\partial[Z]_{\mathrm{sys}}^{-1}}{\partial\Delta_m}=-[Z]_{\mathrm{sys}}^{-1}\frac{\partial[Z]_{\mathrm{sys}}}{\partial\Delta_m}[Z]_{\mathrm{sys}}^{-1} \tag{26}$$

can be used in Eq. (25). As a result, Eq. (25) becomes

$$\frac{\partial\{W\}}{\partial\Delta_m}=-[Z]_{\mathrm{sys}}^{-1}\frac{\partial[Z]_{\mathrm{sys}}}{\partial\Delta_m}[Z]_{\mathrm{sys}}^{-1}\{F^e\}. \tag{27}$$

The only term in $[Z]_{\mathrm{sys}}$ that depends on Δ_m is D_{jk}^m. As such, using Eqs. (22) and (17),

$$\frac{\partial\{W\}}{\partial\Delta_m}=\frac{\rho_s h}{\rho a}(ka)^2[Z]_{\mathrm{sys}}^{-1}\frac{\partial[D^m]}{\partial\Delta_m}[Z]_{\mathrm{sys}}^{-1}\{F^e\}. \tag{28}$$

Since μ_{Jk} and κ_{jk} do not depend on Δ_m, the surface displacement sensitivity can be further simplified as

$$\frac{\partial\{W\}}{\partial\Delta_m}=\frac{\rho_s h}{\rho a}(ka)^2[Z]_{\mathrm{sys}}^{-1}\frac{\partial[\eta]}{\partial\Delta_m}[Z]_{\mathrm{sys}}^{-1}\{F^e\}. \tag{29}$$

The resulting vector contains the sensitivity of the surface displacement coefficients to changes in the inertial scale. This vector, along with the solution vector $\{W\}$, can be examined to see which displacement coefficients are most sensitive to changes in the inertial scale.

Even though in evaluating Eq. (29) it is necessary to invert the $[A]$ matrix [see Eq. (23)] of size $N\times N$, which is usually large, this inversion need only be carried out once at each excitation frequency. Then, to evaluate the sensitivity at different inertial scales, it is only necessary to compute the inverse of $[Z]_{\mathrm{sys}}$ which is of size $M\times M$. For the system considered here, M is much smaller than N.

Equation (29) is an important result in that all of the terms except $\partial[\eta]/\partial\Delta_m$ are already known. However, the derivative of the inertial coupling coefficient can be found directly from Eq. (18) as

$$\frac{\partial\eta_{jk}}{\partial\Delta_m}=\frac{\partial}{\partial\Delta_m}\left(\frac{m_{rat}}{\Delta_m}\right)\int_{\xi_0-\Delta_m/2}^{\xi_0+\Delta_m/2}\phi_j\phi_k\,d\xi$$
$$+\frac{m_{rat}}{\Delta_m}\frac{\partial}{\partial\Delta_m}\left(\int_{\xi_0-\Delta_m/2}^{\xi_0+\Delta_m/2}\phi_j\phi_k\,d\xi\right). \tag{30}$$

The first term in Eq. (30) represents the effect due to the change in mass density, m_{rat}/Δ_m, at a fixed mass width. The second term, conversely, represents the effect due to changing the area of contact between the plate and the mass, or footprint, at a fixed mass density. These terms will hereafter be referred to as the density and footprint terms, respectively. Recall that throughout this analysis the total mass added to the system remains constant. Nevertheless, if it is desired to investigate the effect of changing either the density or the footprint independently, then the appropriate term in Eq. (30) can be removed. However, setting either term in this relation to zero results in a sensitivity relation that includes changes in the total mass of the system. Therefore, at very small distribution widths (i.e., as Δ_m approaches zero) the individual terms become meaningless since small changes in Δ_m result in a very large change to the amount of added mass. When both terms are considered, however, the large terms add to produce a valid sensitivity.

Consider that for fluid-loaded structures, the concept of a mode is meaningless. One may use any suitable basis functions, as we have used here, in a series solution. Now, once the relative amplitudes of the series functions are known, the question arises as to how to interpret the results. Typically, one may transform the series representation into a wave-number space, effectively decomposing the solution into an infinite distribution of traveling waves. Those wave-number components that are supersonic then account for the radiation. Alternatively, one may decompose the solution into a finite set of ''radiation modes,'' each with well-defined radiation properties Typically, only a limited set of these modes contribute in any significant way.

Both representations, wave number and radiation modes, are completely valid alternative methods. Given the preceding formulation, however, it is convenient to implement the radiation mode technique. As such, the radiation modes for the plate system can be found by applying the method pre-

sented by Chen and Ginsberg.[10] Following that work, the eigenvalue problem

$$[h_R]\{\chi\}_j = \lambda_j [h_I]\{\chi\}_j \tag{31}$$

is solved for the eigenvalues, λ_j, and the corresponding eigenvectors, $\{\chi\}_j$. Here

$$[h] = i2\pi ka[R][A]^{-1}[R]^T \quad \text{and} \quad [h] = [h_R] - i[h_I]. \tag{32}$$

The eigenvectors are normalized according to $\{\chi\}_j^T[h_I]\{\chi\}_j = 1$. Arranging the eigenvectors in column order,

$$[\chi] = [\{\chi\}_1 \quad \{\chi\}_2 \quad \cdots], \tag{33}$$

the shape of the frequency-dependent velocity radiation modes, Φ_j, are found from

$$\{\Phi\} = [\chi]^T\{\phi\}. \tag{34}$$

The corresponding radiation efficiency σ_j for each radiation mode is given by

$$\sigma_j = \frac{\lambda_j}{\nu_j} \quad \text{where} \quad \nu_j = \int_S \Phi_j^2 dS. \tag{35}$$

Then, the velocity coefficients $\{\nu\}$, found above in Eq (5), can be transformed into velocity radiation mode coefficients $\{V\}$ by

$$\{V\} = [\chi]^{-1}\{\nu\}. \tag{36}$$

Note that it is only necessary to solve the eigenvalue problem once at each excitation frequency. Consequently, for a given velocity sensitivity

$$\frac{\partial\{\nu\}}{\partial\Delta_m} = -i(ka)\frac{\partial\{W\}}{\partial\Delta_m}, \tag{37}$$

the sensitivity of the velocity radiation mode coefficients can now be determined. Since the matrix of eigenvectors $\{\chi\}$ and the eigenvalues λ_j do not depend on the presence of the attached mass, it follows from Eq. (36) that

$$\frac{\partial\{V\}}{\partial\Delta_m} = [\chi]^{-1}\frac{\partial\{\nu\}}{\partial\Delta_m}. \tag{38}$$

Herein lies the advantage to this method. The velocity radiation mode shapes do not depend on the internal structure of the system. In fact, the mode shapes only depend on the external geometry of the radiator and the frequency of excitation.[9,16] Furthermore, the radiation efficiency for each radiation mode is also independent of the internal structure of the system. Therefore, once the radiation mode shapes are found for a given excitation frequency, changing the scale of the attached mass only affects the resulting coefficients $\{V\}$.

Now, consider the sensitivity of the surface pressure coefficients to changes in the feature scale. Since the $[R]$ and $[A]$ matrices do not depend on Δ_m, a relationship for the surface pressure can be found from Eqs. (6) and (29) as

$$\frac{\partial\{P\}}{\partial\Delta_m} = 2\pi(ka)^2[A]^{-1}[R]^T\frac{\partial\{W\}}{\partial\Delta_m} \tag{39}$$

or

$$\frac{\partial\{P\}}{\partial\Delta_m} = 2\pi\frac{\rho_s h}{\rho a}(ka)^4[A]^{-1}[R]^T[Z]_{\text{sys}}^{-1}$$
$$\times\frac{\partial[\eta]}{\partial\Delta_m}[Z]_{\text{sys}}^{-1}\{F^e\}. \tag{40}$$

Finally, consider the effects of changing the inertial scale on the radiated acoustic power. Since there are no damping mechanisms in the structure, the power input is equal to the power radiated into the fluid. Therefore, the radiated acoustic power can be expressed as

$$\wp_{\text{rad}} = \tfrac{1}{2}\text{Re}(\{\nu^*\}^T\{F^e\}) \quad \text{or} \quad \wp_{\text{rad}} = \tfrac{1}{2}\text{Re}(\{\nu\}^T\{F^e\}^*), \tag{41}$$

where the asterisk denotes the complex conjugate. Although not always the case, the vector $\{F^e\}$ is purely real for the analyses considered here. Therefore, the second form of Eq. (41) will be used for convenience. Using Eqs. (21) and (5) to find $\{\nu\}$, the radiated power can also be expressed as

$$\wp_{\text{rad}} = \frac{ka}{2}\text{Re}(-i\{F^e\}^T[Z]_{\text{sys}}^{-1}\{F^e\}^*). \tag{42}$$

In comparison to the power relationship used by Salagame et al.,[6] it can be seen that the relationship developed here does not explicitly depend on the surface velocity. Furthermore, unlike the impedance matrix used in that work the impedance matrix in Eq. (42) contains the description of the structure. Taking the derivative of the power in the same manner as above,

$$\frac{\partial\wp_{\text{rad}}}{\partial\Delta_m} = \frac{1}{2}\text{Re}\left(-ika\{F^e\}^T\frac{\partial[Z]_{\text{sys}}^{-1}}{\partial\Delta_m}\{F^e\}^*\right)$$
$$+\text{Re}\left(-ika\frac{\partial\{F^e\}^T}{\partial\Delta_m}[Z]_{\text{sys}}^{-1}\{F^e\}^*\right). \tag{43}$$

Since the forcing vector $\{F^e\}$ is not a function of the mass width, the second term is zero. As a result, once the system parameters have been defined it is only necessary to find $\partial[\eta]/\partial\Delta_m$ to determine the sensitivity of the radiated power with respect the inertial scale. Using Eqs. (30) and (27), it follows that Eq. (43) can be rewritten as

$$\frac{\partial\wp_{\text{rad}}}{\partial\Delta_m} = -\frac{ka}{2}\text{Re}\left[i\{F^e\}^T\frac{\partial\{W\}^*}{\partial\Delta_m}\right] \tag{44}$$

or

$$\frac{\partial\wp_{\text{rad}}}{\partial\Delta_m} = -\frac{\rho_s h}{\rho a}\frac{(ka)^3}{2}$$
$$\times\text{Re}\left[i\{F^e\}^T[Z]_{\text{sys}}^{-1}\frac{\partial[\eta]}{\partial\Delta_m}[Z]_{\text{sys}}^{-1}\{F^e\}^*\right]. \tag{45}$$

As can be seen in this and previous relations, all of the effects of changing the feature scale are contained in the term $\partial[\eta]/\partial\Delta_m$. In all of the above sensitivity relationships, it is possible to separate the sensitivity into a density effect and footprint effect using the appropriate terms in Eq. (30). However, as stated for Eq. (30), retaining both terms preserves the total mass of the system.

The relations developed in this section are specifically related to changes in the inertial scale. However, the sensitivity of any of these system variables to changes in any of the structural parameters, such as wave speed, thickness, and plate density, can easily be determined. For example, by replacing the term $\partial[D^m]/\partial\Delta_m$ in relation (28) with $\partial[D^m]/\partial h$, it is possible to investigate the sensitivity of the surface displacement to changes in the plate thickness.

IV. NUMERICAL EXAMPLE

This section presents an example of the sensitivity development of the previous section. First, a description of the physical properties and modeling parameters of the bare plate and the acoustic fluid is given. A description of the substructural feature considered is then presented along with an overview of some results.

The basis functions implemented for the displacement of the plate, Eq. (3), are the orthogonal eigenfunctions for *in vacuo* vibration of a simply supported plate

$$\phi_j = \sin[i\pi(\xi + \tfrac{1}{2})], \quad -\tfrac{1}{2} \leq \xi \leq \tfrac{1}{2} \tag{46}$$

with the parameter a set to a value of unity. Likewise, the basis functions for the surface pressure, Eq. (2), are

$$\psi_k = \sin[j\pi(\xi/\sigma + \tfrac{1}{2})], \quad -\sigma/2 \leq \xi \leq \sigma/2 \tag{47}$$

with the baffle size set to a value of $\sigma = 5$. The properties of the steel plate, which is submerged in seawater,[17] are

$$a/h = 25, \quad \rho_s/\rho = 7.58 \quad c_s/c = 3.84,$$

and the subscript s refers to steel.

Convergence of the numerical method can easily be established by examining the modal amplitudes of the series basis functions for the plate displacement and fluid pressure found using Eqs. (6) and (9). For the examples presented in this section, 20 structural modes, $M = 20$, and 100 fluid modes, $N = 100$, were used for the range of excitation frequencies considered, $0.05 \leq ka \leq 10.0$. The number of modes used is quite adequate as the magnitudes of the higher numbered coefficients in $\{P\}$ and $\{W\}$ are several orders of magnitude smaller than those with lower indices and are declining with increasing index.

The case of a line mass of $m_{rat} = 0.25$ attached to the plate at $\xi_0 = 0.2071$ will be considered. Since the distribution is not allowed to extend past the limits of the plate, the maximum distribution width is about $\Delta_m = 0.58$. The system is harmonically excited by a line force located at $\xi_e = -0.25$. With the mass in a lumped configuration, the first three fluid-loaded resonances for this case occur near $k_1 a = 0.2$, $k_2 a = 1.25$, and $k_3 a = 3.35$. As shown in Fig. 4, the error in the radiated acoustic power predicted using a lumped element as opposed to a distributed element,

$$10 \log(\text{Power}_{dis}/\text{Power}_{lumped}),$$

is small for all inertial scales except near these ka values. These errors are predominantly due to shifts in the resonant frequencies. However, the error near the first resonance is quite small. Consequently, the following sensitivity analysis will only be concerned with the range of frequencies around the second and third resonances. Note that there appear to be frequencies at which the radiated power is independent of the scale of the attached mass. For
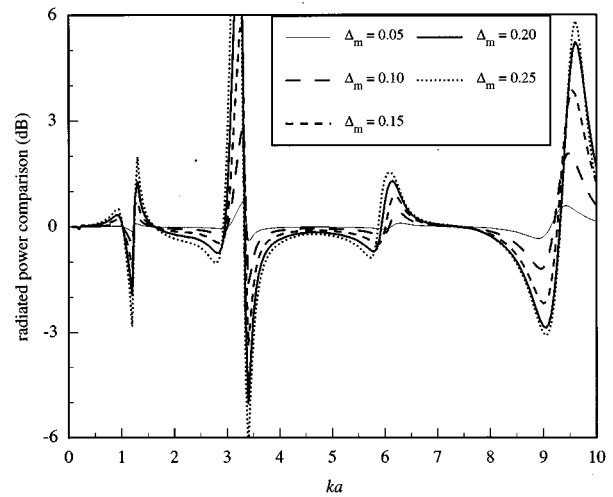
FIG. 4. Comparison of radiated power using a distributed mass of scale Δ_m to power found using a lumped mass. One mass of $m_{rat} = 0.25$ located at $\xi_0 = 0.2071$.

example, near $ka = 1.60$ there is a single frequency at which the power is the same for all Δ_m considered (i.e., 0 dB). These points will be referred to as "zero crossings."

Before considering the sensitivities, it is useful to determine the shape of the velocity radiation modes along with their corresponding radiation efficiencies σ_j. The first four velocity radiation mode shapes are shown in Fig. 5 for $ka = 1.25$. The modes are also represented schematically by the dark lines (see Ref. 9). The mode shapes shown include Gibbs effects due to the limited number of structural modes. It is important to realize that the shapes of the radiation modes do not significantly change for the frequency ranges considered below, $ka \leq 4.0$. The shape of the first mode, for example, remains reasonably constant up to approximately $ka = 5.0$ (see p. 2176, Ref. 16). The radiation efficiencies for these modes are shown in Fig. 6 for the range of frequencies that will be examined. These efficiencies, like the mode shapes, do not depend on the nature of the internal structure.
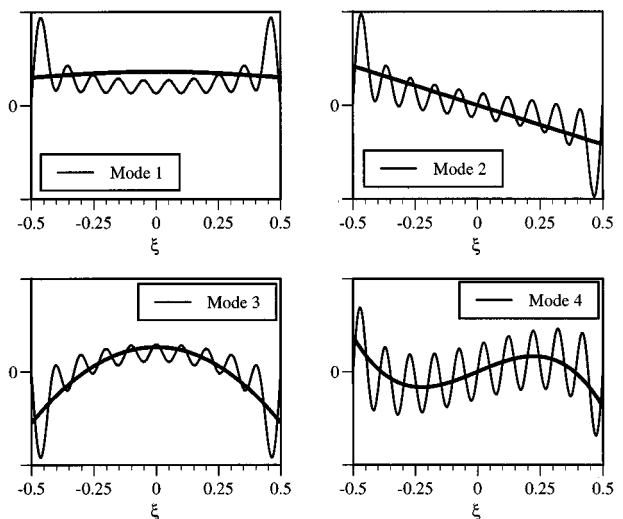
FIG. 5. First four velocity radiation modes for fluid-loaded plate system. $ka = 1.25$.
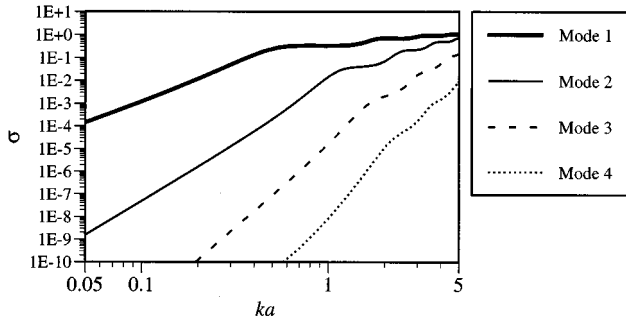
FIG. 6. Radiation efficiencies for first four velocity radiation modes.

The sensitivity of the velocity radiation mode coefficients to changes in inertial scales will now be considered. The magnitude of the first four velocity radiation mode coefficient sensitivities, normalized with respect to the radiation mode coefficients,

$$\left| \left(\frac{\partial V_j}{\partial \Delta_m} \right) \middle/ V_J \right|_{\Delta_{m0}} \qquad j = 1, \ldots, 4 \qquad (48)$$

is shown in Fig. 7 for the frequency range around the second resonance. The sensitivity is evaluated at the scale Δ_{m0}. Although one could argue that an attached mass of scale $\Delta_{m0} = 0.58$ is not usually encountered in most systems, it is nevertheless considered in this research. In these plots, the light regions represent low sensitivity while the dark regions represent high sensitivity. It is interesting to note that the plots for the even modes are similar while the plots for the odd modes are similar. Considering the radiation efficiencies of these modes, however, it becomes evident that only the effects on the first mode will be observed in the far field. Examining just the first mode, there is a small range of frequencies over which it is critical that the feature be modeled accurately to obtain a valid prediction. For example, at $ka = 1.10$ it is more critical to accurately represent a mass of scale near $\Delta_{m0} = 0.10$ than say a mass of $\Delta_{m0} = 0.40$. The converse becomes the case, however, near $ka = 1.20$.

The sensitivities can be broken down into the effects of independently changing the mass density or mass footprint. Consider the sensitivity of the first radiation mode coefficient to changes in mass density only by setting the second term in
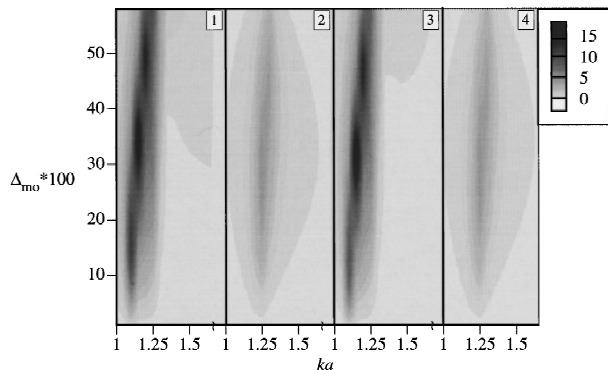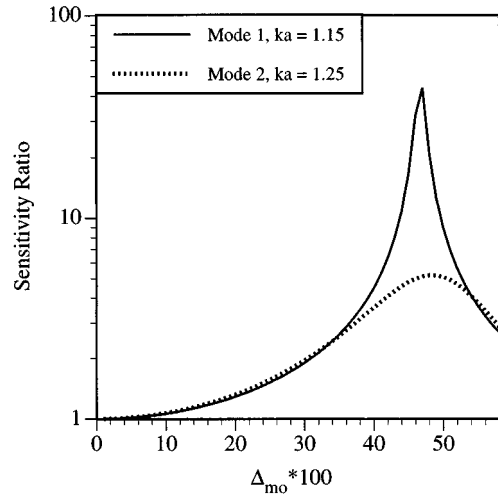
FIG. 8. Sensitivity ratio (see text) for first mode at $ka = 1.15$ and second mode at $ka = 1.25$.

Eq. (30) to zero. Next, consider the sensitivity of the same coefficient to changes in the footprint only by setting the first term of Eq. (30) to zero. The magnitude of the ratio of the density value to the footprint value is hereafter referred to as the sensitivity ratio. This ratio is shown as a function of Δ_{m0} for the first mode at $ka = 1.15$ and the second mode at $ka = 1.25$ in Fig. 8. At low inertial scales, $\Delta_{m0} \lesssim 0.10$, this sensitivity ratio is relatively small. However, as Δ_{m0} increase, the ratio gets quite large. Thus, at large distribution widths the amplitudes of these radiation modes are more sensitive to changes in the mass density than changes in the mass footprint. This makes sense physically since the density is changing over a large portion of the plate. When the footprint is changed at this large scale, however, only a small region of the plate is affected. Note that these sensitivities must be considered together. As stated above, implementing just one term results in a change of the total mass in the system.

For the frequency range around $ka = 3.35$, Fig. 9 shows the normalized magnitude of the coefficient sensitivities. In contrast to the previous frequency range, all of the radiation mode coefficients have similar sensitivity to changes in the feature scale. The first mode, however, is the most sensitive.

FIG. 7. Sensitivity of first four velocity radiation modes coefficients to changes in the inertial scale Δ_m at Δ_{m0}.
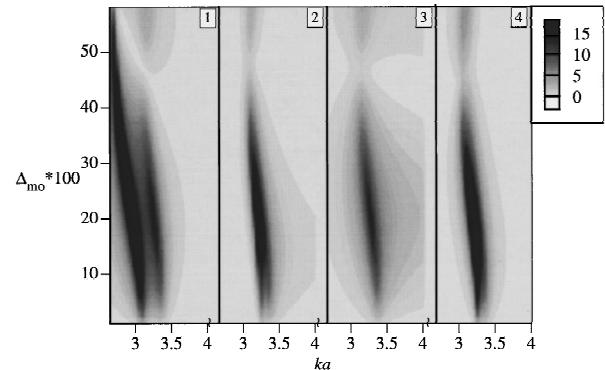
FIG. 9. Sensitivity of first four velocity radiation modes coefficients to changes in the inertial scale Δ_m at Δ_{m0}.
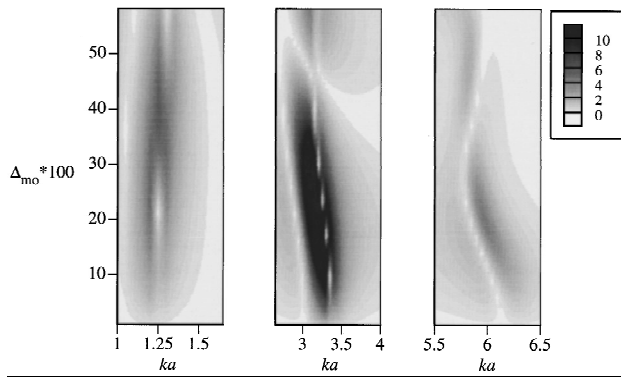
FIG. 10. Magnitude of radiated acoustic power sensitivity to changes in the inertial scale Δ_m at Δ_{m0}.

Furthermore, the radiation efficiency of the first mode is larger than it was in the previous frequency range considered. As a result, the sensitivity of the radiated power should be larger for this frequency range, $2.65 \leqslant ka \leqslant 4.0$, than for the previous range, $1.0 \leqslant ka \leqslant 2.65$. The behavior of the curves in Fig. 4 supports this statement. Note the region of low sensitivity that separates the two regions of high sensitivity for the first mode near $ka = 3.25$. The acoustic response is not sensitive to changes in scale near this ka as also evidenced by the ''zero crossing'' in Fig. 4. At this single frequency, therefore, the scale of the mass becomes transparent to the system response. The range of frequencies where accurate modeling of the attached mass is required, though, is broader than for the preceding frequency range.

The sensitivity of the radiated power for various ranges of excitation frequency is shown in Fig. 10. This is the ratio of the real part of the power sensitivity to the radiated power

$$\left| \frac{\partial \wp_{\text{rad}} / \partial \Delta_m}{\wp_{rad}} \right|_{\Delta_{m0}, ka}. \tag{49}$$

Since this ratio is always a real number, the magnitude is only used to keep all values positive. Even though the radiated power is always positive, the sensitivity can be a negative value. Following from above, it is expected that the radiated power is more sensitive to changes in scale in the second range of frequencies than the first. The lines of very small sensitivity that cross the area of high sensitivity explain the ''zero crossing'' behavior in Fig. 4. The small sensitivity region moves with frequency in the $5.5 \leqslant ka \leqslant 6.5$ frequency range. Thus, the ''zero crossings'' in Fig. 4 for this range are not confined to a small region or point as for the lower frequency range.

V. CONCLUSIONS

We have presented here a new approach to examining attached features in structural acoustics problems: Sensitivity of important system variables to changes in an attached feature's scale. The research tool presented allows close examination of the effects of inertial scales. The effects of changing the scale on the radiated power, surface velocity, and surface pressure were investigated. It was shown that the

relation required determining only one additional term. All of the other terms necessary to evaluate the sensitivities are found during the normal course of modeling the system. The method furthermore used orthogonal velocity radiation modes to decompose the surface velocity coefficient sensitivities into that of mode coefficient sensitivities. These modes have the important and useful property of being independent of the internal structure. The ultimate goal of this continuing research is to make generalized statements concerning the critical scales in structural acoustics systems. This information will then allow structural acoustic modeling and subsequent analyses to be performed in a more efficient manner.

[1] M. M. Ettouney, R. P. Daddazio, and N. Abboud, ''Behavior of submerged shell-structure systems,'' in Proceedings of Winter Annual Meeting, ASME, New Orleans, LA, 1993, 93-WA/NCA-21, pp. 1–9 (unpublished).

[2] D. Feit and J. M. Cuschieri, ''Scattering of sound by a fluid-loaded plate with a distributed mass inhomogeneity,'' J. Acoust. Soc. Am. **99**, 2686–2700 (1996).

[3] B. P. Wang, ''Eigenvalue sensitivity with respect to location of internal stiffness and mass attachments,'' AIAA J. **31**, 791–794 (1992).

[4] R. L. St. Pierre, Jr. and G. H. Koopmann, ''Minimization of radiated sound power from plates using distributed point masses,'' in Proceedings of Winter Annual Meeting, ASME, New Orleans, LA, 1993, 93-WA/NCA-11, pp. 1–9 (unpublished).

[5] K. Naghshineh, G. H. Koopmann, and A. D. Belegundu, ''Material tailoring of structures to achieve a minimum radiation condition,'' J. Acoust. Soc. Am. **92**, 841–855 (1992).

[6] R. R. Salagame, A. D. Belegundu, and G. H. Koopmann, ''Analytical sensitivity of acoustic power radiated from plates,'' J. Vib. Acoust. Trans. ASME **117**, 43–48 (1995).

[7] S. R. Hahn and A. A. Ferri, ''Perturbation methods for sensitivity analysis of structural acoustic models,'' in Proceedings of the Third International Congress on Air- and Structure-Borne Sound and Vibration, Montreal, Canada, 1994, Vol. 3, p. 1799 (unpublished).

[8] S. R. Hahn and A. A. Ferri, ''Sensitivity analysis of coupled structural-acoustic problems using perturbation techniques,'' J. Acoust. Soc. Am. (to be published).

[9] K. A. Cunefare and M. N. Currey, ''On the exterior acoustic radiation modes of structures,'' J. Acoust. Soc. Am. **96**, 2302–2312 (1994).

[10] P. T. Chen and J. H. Ginsberg, ''Complex power, reciprocity, and radiation modes for submerged bodies,'' J. Acoust. Soc. Am. **98**, 3343–3351 (1995).

[11] J. H. Ginsberg, A. D. Pierce, and X. F. Wu, ''A Variational Principle for Sound Radiation from Vibrating Bodies,'' Georgia Institute of Technology Report No. GTADRL-86-101, 1986 (unpublished).

[12] J. H. Ginsberg, K. A. Cunefare, and H. Pham, ''Spectral description of inertial effects in fluid-loaded plates,'' J. Vib. Acoust. Trans. ASME **117**, 206–212 (1995).

[13] J. H. Ginsberg and P. Chu, ''Asymmetric vibration of a heavily fluid-loaded circular plate using variational principles,'' J. Acoust. Soc. Am. **91**, 894–906 (1992).

[14] J. H. Ginsberg and J. G. McDaniel, ''An acoustic variational principle and component mode synthesis applied to the analysis of acoustic radiation form a concentrically stiffened plate,'' J. Vib. Acoust. Trans. ASME **113**, 401–408 (1991).

[15] R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary Matrices and Some Applications to Dynamics and Differential Equations* (Cambridge, U.P. Cambridge, England, 1938), p. 416.

[16] M. E. Johnson and S. J. Elliott, ''Active control of sound radiation using volume velocity cancellation,'' J. Acoust. Soc. Am. **98**, 2174–2186 (1995).

[17] S. Alper and E. B. Magrab, ''Radiation from the forced harmonic vibrations of a clamped circular plate in an acoustic fluid,'' J. Acoust. Soc. Am. **48**, 681–691 (1970).

Power flow in an aeroelastic plate backed by a reverberant cavity

Kenneth D. Frampton and Robert L. Clark

Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708-0302

This paper investigates the modeling of an elastic plate coupled to a reverberant acoustic enclosure and subjected to convected fluid loading. The primary objective of this work is to quantify the effects of variations in external fluid convection velocity on power flow from the plate. The plate power flow constituents consist of injected power due to turbulent boundary layer (TBL) pressure disturbance, sound power radiated to the convected flow, sound power radiated to a reverberant cavity, and power dissipated through structural damping. Results indicate that variations in the convected fluid loading can significantly alter the balance in power flow. These results demonstrate the importance of including the effects of convected fluid loading in modeling the sound transmission through plates. Results also indicate that convected fluid loading serves to decrease the sound power radiated to the cavity as the external flow velocity increases. This effect, however, is overwhelmed by the increase in turbulent boundary layer pressures with the square of the flow speed. This work concludes with a discussion on the potential effectiveness of dissipative control strategies (be they active or passive) toward minimizing the sound transmitted to the cavity. © *1997 Acoustical Society of America.* [S0001-4966(97)01109-0]

PACS numbers: 43.40.Rj, 43.40.Vn [CBB]

LIST OF SYMBOLS

a	plate length (chord)
a_c	cavity speed of sound
a_a	external speed of sound
b	plate width (span)
C_{nk}	cavity/plate coupling coefficient
D	$EI/(1-\nu^2)$, stiffness
D_{mn}	aerodynamic influence coefficient
h	plate thickness
$H(\omega)$	coupled system transfer function
$H^a(\omega)$	aerodynamic system transfer function
$H_{mn}(t)$	aerodynamic influence function
$I_{mn}(t)$	aerodynamic influence function
k,m,n	modal indices
K	number of cavity modes
L_x, L_y, L_z	cavity dimensions
M	Mach number
M_n	modal mass
N	number of plate modes
$p^a(x,y,t)$	aerodynamic pressure
$p^c(x,y,t)$	cavity pressure
$p^{\text{tbl}}(x,y,t)$	turbulent boundary layer pressure
$\mathscr{P}_{Q_m Q_n}(\omega)$	TBL generalized force power spectrum
$\mathscr{P}_{pp}(\omega; \eta, \gamma)$	TBL correlation function
$\mathscr{S}_{pp}(\omega)$	TBL point pressure power spectrum
$q_n(t)$	plate generalized coordinate
$Q_n^a(t)$	aerodynamic generalized force
$Q_n^{tbl}(t)$	TBL generalized force
$r_k(t)$	cavity pressure generalized coordinate
S_{mn}	aerodynamic influence coefficient
t	time
U_a	free stream velocity
U_c	TBL convection velocity
$w(x,y,t)$	panel displacement
x,y,z	Cartesian coordinates
α_1, α_2	TBL model spatial correlation constants
ξ	plate modal damping coefficient
ξ^c	cavity modal damping coefficient
$\Gamma_k(x,y,z)$	cavity modal expansion functions
Λ_k	cavity modal volume
Π_i	TBL input power
Π_d	plate damping dissipated power
Π_f	power radiated to convected fluid
Π_c	power radiated to cavity fluid
$\Psi_n(x,y)$	plate modal expansion functions
$\Phi^a(x,y,z,t)$	external fluid velocity potential
$\phi_k^c(t)$	cavity generalized coordinate
$\Phi^c(x,y,z,t)$	cavity fluid velocity potential
ρ_a	external fluid density
ρ_c	cavity fluid density
ρ_p	plate density
ω_n	plate natural frequency
ω_k^c	cavity natural frequency

INTRODUCTION

The reduction of sound transmission to aircraft interiors is an objective of modern airframe manufacturers. The dominant noise sources incident on aircraft exteriors are engine noise and turbulent boundary layer (TBL) pressures. TBL pressures are widely regarded as the most important noise source for jet powered aircraft at cruise speeds particularly as new, quieter jet engines are developed. In an attempt to iden-

tify effective, weight saving noise control strategies active control of sound transmission through elastic plates has received much attention in the literature. While many of these investigations have demonstrated promising results, none of them have considered the effects on plate dynamics by convected fluid loading. Past investigations into the effects of convected fluid loading on plates (i.e., aeroelasticity) have demonstrated significant effects on plate response in the high subsonic and supersonic Mach number range. This is the typical operating range for current and future commercial aircraft.

The effects of convected fluid loading on elastic plates has received a great deal of attention in the literature. Publications by Dowell,[1–3] Dugundji,[4] and Hedgepeth[5] investigated the parametric stability boundaries of plates subjected to convected fluid loading (i.e., panel flutter). Results from these investigations demonstrate that convected fluid loading can have a large effect on plate dynamics.

A large volume of work is also available concerning reverberant cavity backed plates. Work by Dowell and Voss,[6] Fahy,[7] Pretlove,[8] as well as Guy and Bhattacharya[9] demonstrate modeling methods for cavity backed plates and demonstrate the dynamic effects which were found to be significant.

The effects of external fluid flow on the sound transmission have received only limited attention in the literature. Dowell investigated a convected fluid loaded plate backed by a cavity with an emphasis on plate response.[10] More recently Koval considered the sound transmission through a convected fluid loaded infinite plate.[11] Results from this investigation show slight decreases in sound transmission with increasing flow velocity. Atalla and Nicolas[12] and later Howe and Shah[13] investigated the effects of convection on acoustic radiation from plates. Both of these studies found that as flow increased, the efficiency of radiation to the convected fluid increased. Finally, Clark and Frampton considered sound transmission through a convected fluid loaded finite plate with similar results.[14,15]

Many applications of active control strategies on sound transmission and radiation from plates have been published. Some of the earliest efforts were published by Clark and Fuller[16,17] which applied adaptive feedforward control techniques with piezoelectric sensors and actuators to harmonic disturbances. The feedforward approach utilized by Clark and Fuller was enhanced by Smith et al.[18] to include broadband disturbance control with promising results. A more recent effort was published by Thomas and Nelson[19] which applied optimal state feedback strategies to the problem of turbulent boundary layer noise transmission control. The control system was designed based on a complete model of TBL pressure excitation, plate dynamics, and sound radiation filters. The next step in this evolution is to include the dynamic effects of convected fluid loading on the plate in the control system design.

The objective of this work is to quantify the power flow from a convected fluid loaded plate subjected to turbulent boundary layer noise and coupled to a reverberant acoustic enclosure and to use this analysis to identify fundamental strategies for the active control of noise transmitted to the
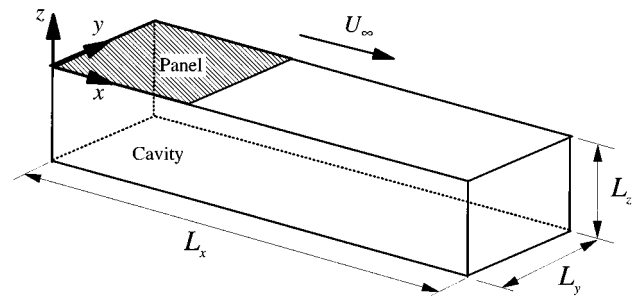
FIG. 1. System schematic.

cavity. This is accomplished by presenting the system modeling technique which is based, in part, on the aeroelastic plate model developed by Frampton et al.[20] and employing a singular value decomposition technique for the aerodynamics as presented by Frampton and Clark.[21] The vibrational power flow from the plate including that to the cavity, to structural losses, and to the convected fluid, will be quantified, and the effects of changes in the flow velocity on power flow will be discussed. Finally, the effectiveness of a dissipative active control strategy will be discussed.

I. MODEL DEVELOPMENT

A diagram of the system is shown in Fig. 1. The plate occupies a portion of the reverberant cavity side wall and is located in an infinite baffle. The plate is subjected to convected fluid loading and turbulent boundary layer (TBL) pressures on the surface opposite the cavity.

The model development consists of four parts: (1) the plate model, (2) the external aerodynamic model, (3) the cavity model, and (4) the TBL noise model.

A. Plate dynamics

The plate is modeled using Galerkin's method to discretize the linear partial differential equation describing the motion of a thin, uniform panel[22] including the internal and external pressure loads

$$D\left[\frac{\partial^4 w(x,y,t)}{\partial x^4} + \frac{\partial^4 w(x,y,t)}{\partial x^2 \partial y^2} + \frac{\partial^4 w(x,y,t)}{\partial y^4}\right]$$
$$+ \rho_p h \frac{\partial^2 w(x,y,t)}{\partial t^2}$$
$$= p^c(x,y,z=0,t) - p^a(x,y,z=0,t) - p^{\text{tbl}}(x,y,z=0,t). \quad (1)$$

This expression assumes that the external pressures which are generated by the panel motion, p^a, and those associated with the turbulent boundary layer disturbance, p^{tbl} are additive. This assumption is valid when the turbulence velocities are much greater than the acoustic velocities.[23] A solution is assumed based on the *in vacuo* plate eigenfunctions and generalized coordinates of the form

$$w(x,y,t)=\sum_{n=1}^{N}\Psi_n(x,y)q_n(t). \qquad (2)$$

Substituting Eq. (2) into Eq. (1), multiplying by an arbitrary expansion function, $\Psi_m(x,y)$, and integrating over the plate surface yields a set of ordinary differential equations of the form:

$$M_n[\ddot{q}_n(t)+2\xi\omega_n\dot{q}_n(t)+\omega_n^2 q_n(t)]+\rho_a U_a^2 Q_n^a(t)+Q_n^{\text{tbl}}(t)$$

$$+\rho_c ab\sum_{k=1}^{K}C_{nk}\dot{\phi}_k^c(t)=0. \qquad (3)$$

where $n=1,2,...,N$. To account for energy dissipation in the plate a proportional modal damping term, $2\xi\omega_n\dot{q}_n$, has been added to the modal equations where ξ is the damping ratio.

The aerodynamic generalized forces are given by

$$Q_n^a(t)=\int_0^b\int_0^a \frac{p^a(x,y,z=0,t)}{\rho_a U_a^2}\Psi_n(x,y)dx\ dy, \qquad (4)$$

and the turbulent boundary layer generalized forces are similarly defined by

$$Q_n^{\text{tbl}}(t)=\int_0^b\int_0^a p^{tbl}(x,y,z=0,t)\Psi_n(x,y)dx\ dy. \qquad (5)$$

The coupling to the reverberant acoustic cavity is expressed by the term $\rho_c ab\Sigma_{k=1}^{K}C_{nk}\dot{\phi}_k^c(t)$ and is detailed in a later section.

For purposes of this investigation the plate is assumed to be simply supported and only modes in the direction of external fluid flow are considered. This assumption was made since convected fluid loading mainly affects these modes. This assumption also greatly reduced the computational burden. Thus, the plate expansion functions are of the form

$$\Psi_n(x,y)=\sin\left(\frac{n\pi}{a}x\right)\sin\left(\frac{\pi}{b}y\right). \qquad (6)$$

B. External flow generalized forces

The generalized forces on the plate due to the external flow, $Q_n^a(t)$, are found by solving the convected wave equation describing the velocity potential in an inviscid, irrotational fluid moving parallel to the x axis,

$$\nabla^2\Phi^a-\frac{1}{a_a^2}\left[\frac{\partial^2\Phi^a}{\partial t^2}+2U_a\frac{\partial^2\Phi^a}{\partial x\ \partial t}+U_a^2\frac{\partial^2\Phi^a}{\partial x^2}\right]=0, \qquad (7)$$

subject to the boundary conditions for a plate in an infinite baffle

$$\left.\frac{\partial\Phi^a}{\partial z}\right|_{z=0}=\begin{cases}\dfrac{\partial w}{\partial t}+U_a\dfrac{\partial w}{\partial x} & \text{on the panel}\\[2mm] 0 & \text{off the panel}\end{cases} \qquad (8)$$

and a finiteness or radiation condition as z approaches infinity.

The solution can be found by taking a Laplace transform with respect to time and a double Fourier transform with respect to the x and y spatial coordinates. The transformation is applied to Eqs. (7) and (8) and the pressure and velocity potential are related by Bernoulli's equation:

$$p^a=-\rho_a\left[\frac{\partial\Phi^a}{\partial t}+U_a\frac{\partial\Phi^a}{\partial x}\right]. \qquad (9)$$

The solution yields the generalized aerodynamic forces on the panel which are given here as

$$Q_n^a(t)=\sum_{m=1}^{N}Q_{mn}^a, \qquad (10)$$

where

$$Q_{mn}^a(t)=q_m(t)S_{mn}+\dot{q}_m(t)D_{mn}+\int_0^t q_m(\tau)H_{mn}(t-\tau)d\tau$$

$$+\int_0^t\dot{q}_m(\tau)I_{mn}(t-\tau)d\tau. \qquad (11)$$

This solution is detailed in several references.[1,3,14,20] The aerodynamic influence functions $H_{mn}(t)$ and $I_{mn}(t)$ are defined by integrals with no closed form solution. Therefore a singular value decomposition technique was used to approximate the aerodynamic system as presented by Frampton and Clark.[21] This technique yields a state variable representation of the aerodynamic coupling to the plate. The resulting system receives the plate generalized coordinates and their rates as inputs and provides the aerodynamic generalized forces on the plate as outputs. This aerodynamic model can then be coupled to the plate model as feedback resulting in a complete aeroelastic model.[20]

C. Acoustic cavity model

The cavity fluid obeys the wave equation such that

$$\nabla^2 p^c(x,y,z,t)-\frac{1}{a_c^2}\frac{\partial^2 p^c(x,y,z,t)}{\partial t^2}$$

$$=2\rho_c\frac{\partial^2 w(x,y,t)}{\partial t^2}\delta(z), \qquad (12)$$

where the spatial Dirac delta function, $\delta(z)$, is used to describe the plate coupling to the cavity. The plate is assumed to act as a set of infinitesimal volume velocity sources in the $z=0$ plane.[24] Expanding the pressure in a series of the hard walled cavity eigenfunctions (corresponding to eigenvalues ω_m^c) such that

$$p^c(x,y,z,t)=\sum_{k=1}^{K}\Gamma_k(x,y,z)r_k(t), \qquad (13)$$

substituting this into Eq. (12), multiplying by an arbitrary expansion function, $\Gamma_n(x,y,z)$, and integrating over the volume yields the equations of motion for the cavity[24]

$$\Lambda_k\ddot{r}_k(t)+\Lambda_k(\omega_k^c)^2 r_k(t)=-\rho_c a_c^2 ab\sum_{n=1}^{N}C_{nk}\ddot{q}_n(t). \qquad (14)$$

where the modal volumes Λ_k are expressed as

$$\Lambda_k = \int_V \Gamma_k^2(x,y,z) dV. \tag{15}$$

For a plate having surface area S the coupling matrix, C_{nk}, is defined as

$$C_{nk} = \frac{1}{S} \int_S \Psi_n(x,y) \Gamma_k(x,y,z=0) dS. \tag{16}$$

It is more convenient to the chosen state variable modeling process to express the cavity dynamics in terms of the velocity potential instead of the pressure.[25] Toward this end, the cavity fluid velocity potential, $\Phi^c(x,y,z,t)$, is related to the cavity pressure such that

$$p^c(x,y,z,t) = -\rho_c \frac{\partial \Phi^c(x,y,z,t)}{\partial t} \tag{17}$$

and can also be expanded over the hard walled cavity functions as follows

$$\Phi^c(x,y,z,t) = \sum_{k=1}^{K} \Gamma_k(x,y,z) \phi_k^c(t), \tag{18}$$

where the generalized coordinate, $\phi_k^c(t)$, is associated with the kth acoustic mode.

The reverberant cavity equation of motion, including a modal damping term, can then be expressed as

$$\Lambda_k \ddot{\phi}_k^c(t) + 2\xi^c \omega_k^c \dot{\phi}_k^c(t) + \Lambda_k(\omega_k^c)^2 \phi_k^c(t)$$
$$= a_c^2 ab \sum_{n=1}^{N} C_{nk} \dot{q}_n(t), \tag{19}$$

where the proportional damping term, $2\xi^c \omega_k^c \dot{\phi}_k^c(t)$, represents a dissipative mechanism in the cavity.

D. Turbulent boundary layer model

The system excitation is turbulent boundary layer (TBL) pressures on the plate exterior. Beginning with the equation for the TBL generalized forces on the plate stated in Eq. (5) and employing fundamental Fourier transform and power spectrum relations,[22] the power spectrum for the generalized TBL forces on the plate are

$$\mathscr{P}_{Q_m Q_n}(\omega) = \int_S \int_S \mathscr{P}_{pp}(\omega; \eta, \gamma) \Psi_m(x,y)$$
$$\times \Psi_n(x^*,y^*) dS \, dS^*. \tag{20}$$

The form of the TBL spatial/temporal correlation function, $\mathscr{P}_{pp}(\omega; \eta, \gamma)$, used here is that presented by Corcos[26] which has the form

$$\mathscr{P}_{pp}(\omega; \eta, \gamma) = \mathscr{S}_{pp}(\omega) \exp\left[-\frac{\omega \eta}{\alpha_1 U_c}\right]$$
$$\times \exp\left[-\frac{\omega \gamma}{\alpha_2 U_c}\right] \exp\left[-\frac{i \omega \eta}{U_c}\right], \tag{21}$$

where $\eta = \|x - x^*\|$ and $\gamma = \|y - y^*\|$ are the distances between correlation points. For the frequency range of interest ($f < 1000$ Hz), the frequency dependent function, $\mathscr{S}_{pp}(\omega)$, is taken to be a constant such that

$$\mathscr{S}_{pp}(\omega) = 3.84 \times 10^{-5} \frac{(\rho_a U_a^2)^2}{4\omega_{\max}} \tag{22}$$

where $\omega_{\max}$ is the maximum frequency of interest.[27] The spatial correlation is controlled by constants α_1 and α_2. These have been empirically determined to be approximately 10 and 2, respectively.[19] The boundary layer convection velocity is assumed to be $U_c = 0.6 U_a$.

A further assumption used is that the cross modal power spectra equals zero, i.e., $\mathscr{P}_{Q_m Q_n} = 0$ for $m \neq n$. This is equivalent to neglecting the modal coupling of the plate resulting from the TBL model. While this is not an accurate assumption for independent TBL disturbances, the TBL is not independent but is accompanied by external flow aerodynamic loading which incorporates the plate modal coupling.

While it is known that this TBL model is not accurate over the complete Mach number range used here, it is accurate in the region of the convective ridge.[28] The plate modes in this model are in the convective ridge region in the mid-subsonic Mach number region ($0.3 < M < 0.6$). The Corcos model does capture the fundamental TBL pressure tendency to roll off in frequency and to scale in amplitude with Mach number squared. The advantage of the Corcos model is that it has a closed form solution for the generalized force power spectra, $\mathscr{P}_{Q_m Q_n}(\omega)$. This results in a significantly reduced computational effort. This work does not focus on advancing TBL noise models, but rather in building a tractable system model which permits the investigation of convected fluid loading effects on sound transmission. More sophisticated TBL models are available which are valid over a greater Mach number range, however, these methods require much more computational effort to employ.[28]

II. POWER CALCULATIONS

The power balance equation can be obtained by multiplying Eq. (3) by the generalized coordinate rate, $\dot{q}_n$, taking the time average and summing over all plate modes[24]

$$0 = \sum_{n=1}^{N} M_n \overline{\ddot{q}_n \dot{q}_n} + \sum_{n=1}^{N} 2\xi M_n \omega_n \overline{\dot{q}_n \dot{q}_n} + \sum_{n=1}^{N} M_n \omega_n^2 \overline{q_n \dot{q}_n}$$
$$+ \rho_a U_a^2 \sum_{n=1}^{N} \overline{Q_n^a \dot{q}_n} + \sum_{n=1}^{N} \overline{Q_n^{tbl} \dot{q}_n}$$
$$+ \rho_c ab \sum_{n=1}^{N} \sum_{k=1}^{K} C_{nk} \overline{\dot{\phi}_k^c \dot{q}_n}, \tag{23}$$

where the time average is defined by

$$\overline{\dot{q}_n \dot{q}_n} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \dot{q}_n(t) \dot{q}_n(t) dt. \tag{24}$$

Note that for time stationary motion the terms $\overline{\ddot{q}_n \dot{q}_n}$ and $\overline{q_n \dot{q}_n}$ are zero due to orthogonality. The second term in Eq. (23) represents the power dissipated through damping in the plate. The fourth term accounts for power radiated to the convected fluid medium. The fifth term is the power input to the system by the TBL pressures and the final term is the

power radiated to (and subsequently dissipated in) the reverberant cavity.

By employing correlation function and Fourier transform relationships the power balance equation can be represented in the frequency domain as

$$\Pi_i(\omega) = \Pi_f(\omega) + \Pi_d(\omega) + \Pi_c(\omega) \qquad (25)$$

where

$$\Pi_i(\omega) = \sum_{n=1}^{N} H_{\dot{q}_n Q_n}(-\omega)\mathscr{P}_{Q_n Q_n}(\omega), \qquad (26)$$

$$\Pi_f(\omega) = \sum_{n=1}^{N}\sum_{k=1}^{N} H^a_{Q_n q_k}(-\omega)H_{q_k Q_n}(-\omega)$$

$$\times H_{\dot{q}_n Q_n}(\omega)\mathscr{P}_{Q_n Q_n}(\omega) + \sum_{n=1}^{N}\sum_{k=1}^{N} H^a_{Q_n \dot{q}_k}(-\omega)$$

$$\times H_{q_k Q_n}(-\omega)H_{\dot{q}_n Q_n}(\omega)\mathscr{P}_{Q_n Q_n}(\omega), \qquad (27)$$

$$\Pi_d(\omega) = \sum_{n=1}^{N} 2\xi\omega_n H_{\dot{q}_n Q_n}(\omega)H_{\dot{q}_n Q_n}(-\omega)\mathscr{P}_{Q_n Q_n}(\omega), \qquad (28)$$

$$\Pi_c(\omega) = \rho_c a b \sum_{n=1}^{N}\sum_{k=1}^{K} C_{nk} H_{\dot{q}_n Q_n}(\omega)$$

$$\times H_{\dot{\phi}_k Q_n}(-\omega)\mathscr{P}_{Q_n Q_n}(\omega). \qquad (29)$$

Here, $H_{q_i Q_j}(\omega)$ refers to the coupled system transfer function between the jth generalized force input and the ith generalized coordinate (or generalized coordinate rate) output. Expressions of the form $H(-\omega)$ refer to the complex conjugate of the transfer function. Calculation also requires the transfer functions for the aerodynamic system alone. The aerodynamic transfer functions of the form $H^a_{Q_j q_i}(\omega)$ refers to the transfer function between the jth generalized coordinate (or generalized coordinate rate) and the ith generalized force. Note that for the aerodynamic system the generalized coordinates are the inputs and the generalized forces are the outputs.[21]

The transfer functions were computed using MATLAB based on the dynamic systems described previously and cast in state variable form.[25] The total power contributions can be found by integrating Eqs. (26)–(29) over all frequencies.

III. RESULTS

Since the primary objective of this work is to quantify the power flow from a convected fluid loaded plate backed by a reverberant cavity and subjected to TBL pressures, the individual power constituents expressed in Eq. (25) were computed for various external flow velocities. In these calculations the following system parameters were held fixed: $L_x = 3.0$ m, $L_y = 0.3$ m, $L_z = 0.3$ m, $a = b = 0.3$ m, $h_s = 1.8$ mm, $D = 35.9$ Nm, $\rho_c = \rho_a = 0.42$ kg/m^3, $a_c = a_a = 310$ m/s. The external flow velocity was varied between 62 and 620 m/s corresponding to Mach numbers of 0.1 to 2.0. These parameters are typical for an aluminum plate at 10-km altitude. The plate is located in the xy plane at one end of the
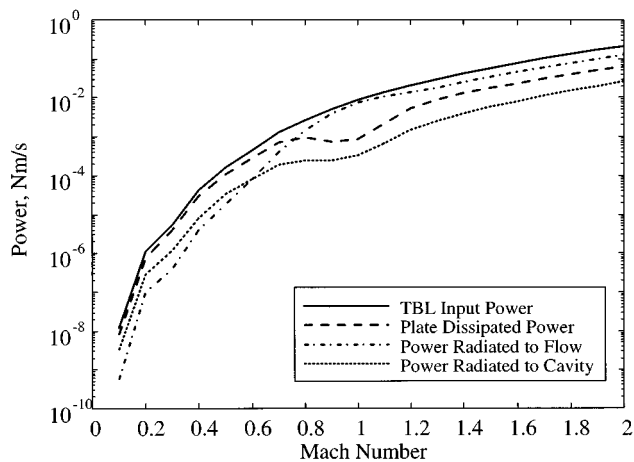
FIG. 2. Power balance versus Mach number.

cavity similar to that shown in Fig. 1. In addition, 4 plate modes and 17 cavity modes were used in the model. The *in vacuo* plate mode resonances occur at 95, 237, 474, and 806 Hz. The *in vacuo* cavity mode resonances occur at 0, 52, 103, 155, 206, 258, 310, 361, 413, 465, 516, 568, 620, 671, 723 775, and 826 Hz. Although an accurate model of a plate in a typical aircraft fuselage would require many more modes, this model should suffice to highlight the physical effects.

The first result is shown in Fig. 2 which depicts each total power constituent with varying Mach number. The total power was found by integrating the power spectrum over a 1000-Hz bandwidth. The trend is for each of the power constituents to increase with increasing Mach number. This is dominated by the TBL pressure amplitude increasing with the square of the Mach number (hence TBL power spectra increases like Mach number to the fourth).

Since it is difficult to see the effects of varying flow speed on the power flow constituents due to the overwhelming increase in TBL input power with increasing Mach number, the three dissipative power constituents (damping, radiation to flow, and radiation to cavity) were normalized with respect to the TBL input power and are shown in Fig. 3. Viewing the dissipative power constituents when normalized
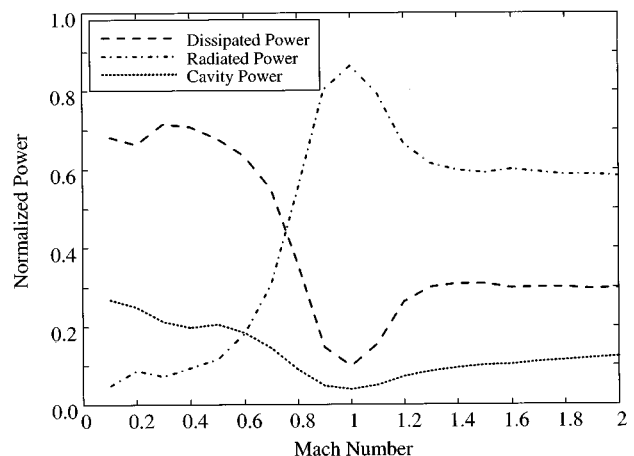
FIG. 3. Normalized power balance versus Mach number.

with respect to the TBL input power is equivalent to holding the TBL input power constant throughout the Mach number range. This serves to independently highlight the effects of increasing external flow velocity on the power flows. Figure 3 shows that the dominant path for energy dissipation changes with increasing Mach number. Most energy is lost to structural damping for Mach numbers below 0.8. At $M = 0.8$ the power loss to structural damping and radiation to the flow field are equal. Above $M = 0.8$ most power is lost as acoustic radiation to the external flow (often referred to as aerodynamic damping[29]).

The effect of increasing Mach number on the power flow to the cavity is also shown in Fig. 3. Note that as flow speed increases the power transmitted to the cavity decreases, particularly near $M = 1.0$. This result is consistent with previously published results.[12,14,23] The decrease in sound transmission can be attributed to an increase in what has previously been termed aerodynamic damping.[29] Aerodynamic damping can be viewed as an increase in the sound power radiated to the convected fluid with increasing flow speed as shown in Fig. 3. While the aerodynamic damping increases, the convected fluid loading also modifies the plate stiffness resulting in an overall change in the plate impedance. In fact, the natural frequency of the fundamental plate mode decreases by more than 40% at $M = 1.0$ relative to the *in vacuo* natural frequency. This is because the convected fluid dynamic pressure is a significant fraction of the divergence dynamic pressure near $M = 1.0$.[3] Divergence is a static instability associated with the convected fluid loading. As the dynamic pressure approaches the divergence dynamic pressure, the fundamental plate mode frequency approaches zero. This movement in the system natural frequencies would be particularly important when designing an active control system. Any control system designed without consideration of these effects would certainly perform nonoptimally in the presence of flow and potentially result in instability.

It is important to remember that the previous discussion focuses on Fig. 3 which shows normalized power. The actual cavity power flow is shown in Fig. 2 which demonstrates that the decrease in sound power transmission to the cavity is overwhelmed by the increase in TBL pressure amplitude with increasing Mach number. Therefore although the increase in flow velocity results in a favorable increase in transmission loss, the scaling of TBL pressures with Mach number results in dramatically increasing power flow to the cavity.

Figure 3 also demonstrates the importance of considering the effects of convected fluid loading when modeling sound transmission. Most studies on TBL sound transmission neglect this aspect of the model.[19,23] However, note that convected fluid loading modifies not only the balance of power flow from the plate but it also significantly modifies the resonance frequencies of the plate. This is demonstrated in the following discussion.

The power spectra of the power flow to the cavity are shown in Figs. 4 and 5 for subsonic and supersonic flow, respectively. At low Mach numbers the TBL pressure amplitude is relatively small resulting in little power flow to the cavity. The magnitude of cavity power flow increases by six
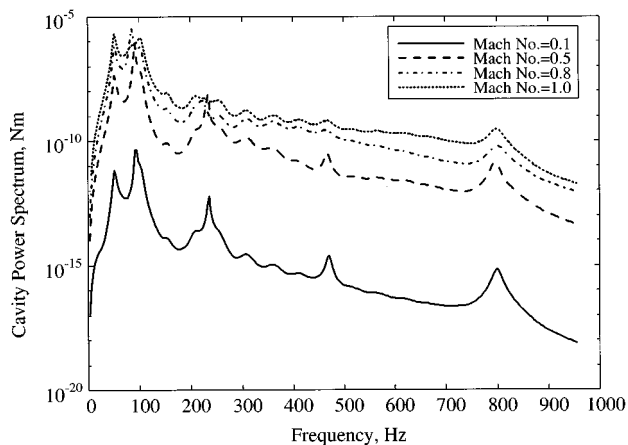
FIG. 4. Cavity power spectra at various subsonic Mach numbers.

orders of magnitude in the subsonic flow range as shown in Fig. 4. For supersonic flow speeds the cavity power continues to increase. Note how the cavity power spectrum tends to roll off with increasing frequency. This is due to the roll off in the TBL power spectrum.

Note in Figs. 4 and 5 that the power radiated to the cavity is a maximum near the second cavity mode rsonance (52 Hz) and the first plate and third cavity mode resonances (97 and 103 Hz, respectively) throughout the Mach number range. This is partly due to the TBL power spectrum being a maximum near these frequencies and partly due to the relatively strong coupling between the fundamental plate mode and the lowest cavity modes. Also note in Figs. 4 and 5 the dramatic shift in frequency and damping of the fundamental plate mode near 100 Hz. This is a result of the convected fluid coupling mentioned previously. These effects are treated more thoroughly by Frampton and Clark.[15]

The dominance of the cavity power flow near the fundamental plate mode indicates that a successful noise transmission control strategy should target this mode. The effectiveness of a dissipative active control approach was estimated by increasing the fundamental plate mode damping ratio from 1% to 10%. The transmission loss across the plate for each case is shown in Fig. 6. Note that increasing the plate fundamental mode damping results in a transmission loss
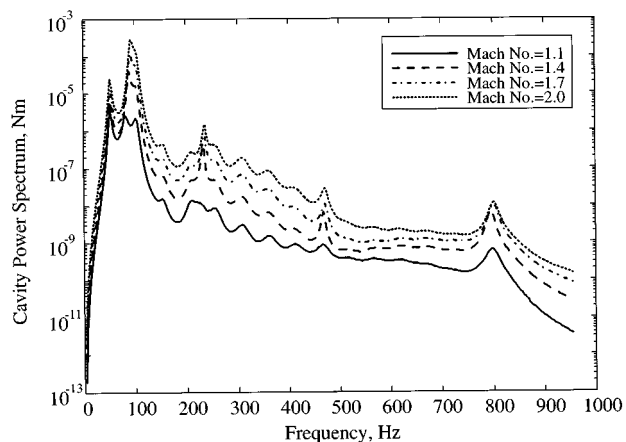
FIG. 5. Cavity power spectra at various supersonic Mach numbers.
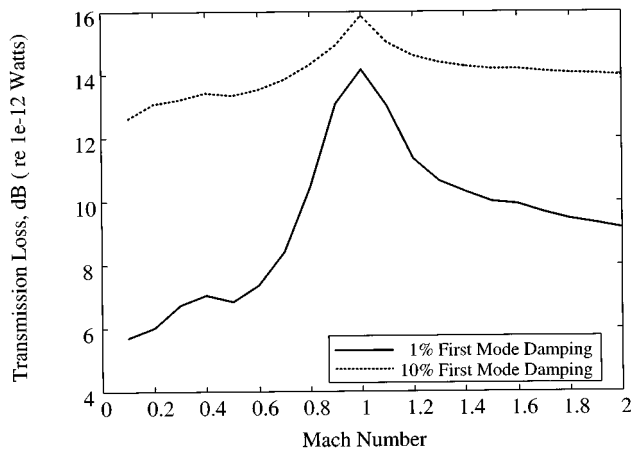
FIG. 6. Transmission loss versus Mach number.

increase of about 7 dB for low Mach numbers. This transmission loss increase is only about 2 dB at $M = 1.0$ increasing again to about 5 dB for $M = 2.0$. This indicates that an active, dissipative control strategy would only be effective for Mach numbers away from $M = 1.0$.

IV. CONCLUSIONS

A method for modeling a plate subjected to convected fluid loading on one side and a reverberant acoustic cavity on the other side has been presented. The effects of convected fluid loading on the transmission of TBL pressures to the cavity has been demonstrated. The primary metric used in this study was the power balance in the plate. The only input power to the system was that due to TBL pressures on the plate. The dissipative power constituents included power loss through structural damping, sound power radiated to the convected fluid, and sound power radiated to the cavity. Results from this study indicate that the sound power radiated to the cavity increases dramatically with increasing Mach number. It was also demonstrated that this trend resulted from TBL pressures increasing with the Mach number squared. When the dissipative power constituents were normalized with respect to the TBL input power it was shown that the effect of mean flow actually decreased the power radiated to the cavity. This was balanced by an increase in power radiated to the convected fluid. The greatest effect was noted around $M = 1.0$.

Results from this study emphasize the importance of including convected fluid loading (or aeroelasticity) in models used to predict the sound transmission to aircraft interiors. The effects of convected fluid loading on the system dynamics are significant, particularly in the high subsonic and supersonic Mach number range. These effects will be particularly important when designing an active structural acoustic control system.

Finally, an estimation was made as to the effectiveness a dissipative active control strategy would have. This was accomplished by increasing the fundamental plate mode damping from 1% to 10%. Results show that this approach will be successful in both subsonic and supersonic flow regimes, but

not near $M = 1.0$. Future investigations will pursue the implementation of robust control strategies to this problem.

[1] E. H. Dowell, ''Generalized aerodynamic forces on a flexible plate undergoing transient motion,'' Q. Appl. Math. **26**, 2267–2270 (1967).
[2] E. H. Dowell, ''Panel flutter: A review of the aeroelastic stability of plates and shells,'' AIAA J. **8**, 385–399 (1970).
[3] E. H. Dowell, *Aeroelasticity of Plates and Shells* (Noordhoff, The Netherlands, 1975).
[4] J. Dugundji, ''Theoretical considerations of panel flutter at high supersonic mach numbers,'' AIAA J. **4**, 1257–1266 (1966).
[5] J. M. Hedgepeth, ''Flutter of rectangular simply supported panels at high supersonic speeds,'' J. Aeronaut. Sci. **24**, 563–573 (1957).
[6] E. H. Dowell and H. M. Voss, ''The effect of a cavity on panel vibration,'' AIAA J. **1**, 476–477 (1962).
[7] F. J. Fahy, ''Vibration of containing structures by sound in the contained fluid,'' J. Sound Vib. **10**, 490–512 (1969).
[8] A. J. Pretlove, ''Forced vibrations of a rectangular panel backed by a closed rectangular cavity,'' J. Sound Vib. **3**, 252–261 (1966).
[9] R. W. Guy and M. C. Bhattacharya, ''The transmission of sound through a cavity-backed finite plate,'' J. Sound Vib. **27**, 207–223 (1973).
[10] E. H. Dowell, ''Transmission of noise from a turbulent boundary layer through a flexible plate into a closed cavity,'' J. Acoust. Soc. Am. **46**, 238–252 (1969).
[11] L. R. Koval, ''Effect of air flow, panel curvature, and internal pressurization on field-incidence transmission loss,'' J. Acoust. Soc. Am. **59**, 1379–1385 (1976).
[12] N. Atalla and J. Nicolas, ''A formulation for mean flow effects on sound radiation from rectangular baffled plates with arbitrary boundary conditions,'' J. Vib. Acoust. **117**, 22–29 (1995).
[13] M. S. Howe and P. L. Shah, ''Influence of mean flow on boundary layer generated interior noise,'' J. Acoust. Soc. Am. **99**, 3401–3411 (1996).
[14] R. L. Clark and K. D. Frampton, ''Aeroelastic structural acoustic coupling: Implications on the control of turbulent boundary layer noise transmission,'' J. Acoust. Soc. Am. (to be published).
[15] K. D. Frampton and R. L. Clark, ''Sound transmission through an aeroelastic plate into a cavity,'' AIAA J. **35,** 1113–1118 (1997).
[16] R. L. Clark and C. R. Fuller, ''Control of sound radiation with adaptive structures,'' J. Intell. Mater. Syst. Struct. **2**, 431–452 (1991).
[17] R. L. Clark and C. R. Fuller, ''Active structural acoustic control with adaptive structures and wavenumber considerations,'' J. Intell. Mater. Syst. Struct. **3**, 296–315 (1992).
[18] J. P. Smith, C. R. Fuller, and R. A. Burdisso, ''Control of broadband acoustic radiation with adaptive structures,'' J. Intell. Mater. Syst. Struct. **7**, 54–64 (1996).
[19] D. R. Thomas and P. A. Nelson, ''Feedback control of sound radiation from a plate excited by a turbulent boundary layer,'' J. Acoust. Soc. Am. **98**, 2651–2662 (1995).
[20] K. D. Frampton, R. L. Clark, and E. H. Dowell, ''State space modeling for aeroelastic panels subject to full potential flow aerodynamic loading,'' J. Aircraft1 **33**, 816–822 (1996).
[21] K. D. Frampton and R. L. Clark, ''State space modeling of aerodynamic forces on a plate using singular value decomposition,'' AIAA J. **34**, 2627–2630 (1996).
[22] L. Meirovitch, *Analytical Methods in Vibrations* (Macmillan, New York, 1967).
[23] W. R. Graham, ''Boundary layer induced noise in aircraft, part i: The flat plate model,'' J. Sound Vib. **192**, 101–120 (1996).
[24] F. Fahy, *Sound and Structural Vibration* (Academic, London, 1985).
[25] K. D. Frampton, ''Active Control of Noise Transmission through an Aeroelastic Plate into an Acoustic Enclosure,'' Ph.D. thesis, Duke University, 1996.

[26] G. M. Corcos, ''The resolution of turbulent pressures at the wall of a boundary layer,'' J. Sound Vib. **6**, 59–70 (1967).

[27] W. K. Blake, ''Turbulent boundary-layer wall-pressure fluctuations on smooth and rough walls,'' J. Fluid Mech. **44**, 637–660 (1970).

[28] M. K. Bull, ''Wall pressure fluctuations beneath turbulent boundary layers: Some reflections on forty years of research,'' J. Sound Vib. **190**, 299–315 (1996).

[29] K. H. Lyle and E. H. Dowell, ''Acoustic radiation damping of flat rectangular plates subjected to subsonic flows. parts i and ii,'' J. Fluids Struct. **8**, 711–746 (1994).

Spatial approximation of leaky wave surface amplitudes for three-dimensional high-frequency scattering: Fresnel patches and application to edge-excited and regular helical waves on cylinders

Philip L. Marston
Department of Physics, Washington State University, Pullman, Washington 99164-2814

For sonar imaging systems and other situations where scattering amplitudes are resolved spatially (e.g., ultrasonic microscopy and nondestructive testing) approximations of outgoing leaky wave amplitudes are needed as a function of position on the imaged surface. The approach developed here approximates the amplitude of a leaky wave pole contribution to the total scattering as a spatial convolution of the local incident pressure with a spatial response function. Leaky rays to a surface point of interest follow a Fermat path having a stationary phase whereas the pole contribution becomes a surface integral that includes defective paths. Increased curvature of the surface or of the incident wavefront ordinarily cause more rapid dephasing along defective paths and a corresponding reduction in size of the Fresnel coupling patch. Examples given include leaky wave excitation on a partially coated cylinder at normal incidence and regular helical leaky wave excitation on tilted cylinders. A helical wave is found to be excited by diffraction at the edge of an idealized coating truncated along the cylinder's axial direction. The leaky wave amplitude becomes proportional to a Fresnel integral of complex argument which accounts for the partial blockage of the Fresnel coupling patch. © *1997 Acoustical Society of America.* [S0001-4966(97)04709-7]

PACS numbers: 43.40.Rj, 43.20.Fn, 43.20.Px, 43.30.Vh [CBB]

INTRODUCTION AND MOTIVATION

With the advent of high-frequency imaging sonar systems (see, e.g., Refs. 1 and 2), there is a need for simple approximations for leaky wave contributions to scattering amplitudes which may be evaluated in the spatial domain of the scatterer. Figure 1 shows an example of a spatially dependent leaky wave scattering process demonstrated in Ref. 1 to greatly enhance the visibility of the ends of an elastic cylinder when the acoustic illumination occurs in the vicinity of certain tilt angles of the cylinder. There can be complications in the modeling of such scattering amplitudes with existing ray formulations. High-frequency sonar systems are typically operated near or above the coincidence frequency for shell-like scatterers of interest, or the scatterers may be natural or man-made solid objects where shell theory is not applicable. Several previous ray formulations for scattering by cylinders, however, use approximations of the mechanical response of the scatterer based on thin shell theories.[3–8] Experimental and theoretical investigations of ray contributions extending far above the coincidence frequency for shells having canonical shapes[9–15] as well as other solid elastic scatterers (e.g., Ref. 16), have revealed features of leaky wave contributions which may be useful in modeling scattering processes for noncanonical geometries as noted in Refs. 17 and 18. Such investigations have examined how leaky wave coupling with the acoustic field varies with frequency and curvature according to elasticity theory.

The purpose of this paper is to present a simple three-dimensional formulation for approximating radiated leaky wave amplitudes. In a subsequent paper,[19] the method is applied to the meridional ray problem shown in Fig. 1 and to other leaky wave contributions for cylinders. The scope of the present paper is to give the formulation along with some elementary (yet nontrivial) examples. Since a spatial formulation is needed, the approach is to generalize the spatial convolution formulation of two-dimensional scattering (Ref. 18, hereafter denoted by M) by embedding the previous formulation in three dimensions. This was chosen for investigation instead of generalizing the formulations in Refs. 3–8 to thick shells and frequencies above coincidence. Some of the previous formulations[3–5] rely on spectral (angular or spatial frequency) calculations while the approach given here directly approximates the spatial evolution of the leaky wave amplitude at the surface of the scatterer given the spatial properties of the illuminating sound which may be a beam of sound. For situations where the illumination either diverges from a source or is focused, the center of divergence or the focal point is assumed to be many wavelengths distant from the surface.

The spatial formulation given here makes it possible to generalize the discussion of the size and shape of the Fresnel coupling patch for the excitation of leaky waves previously given for spheres and cylinders.[17] The identification of Fresnel zones associated with a given ray often gives insight into scattering processes in other situations.[20–22] The Fresnel coupling patch describes the dominant spatial region for the excitation of a given leaky ray; identification of that patch makes it possible to estimate whether modification of a scatterer will have a significant effect on a given leaky ray amplitude. (For example Conti[23] applied the Fresnel patch concept as an aid in understanding leaky wave excitation on a noncanonical scatterer.) Section II discusses the Fresnel cou-
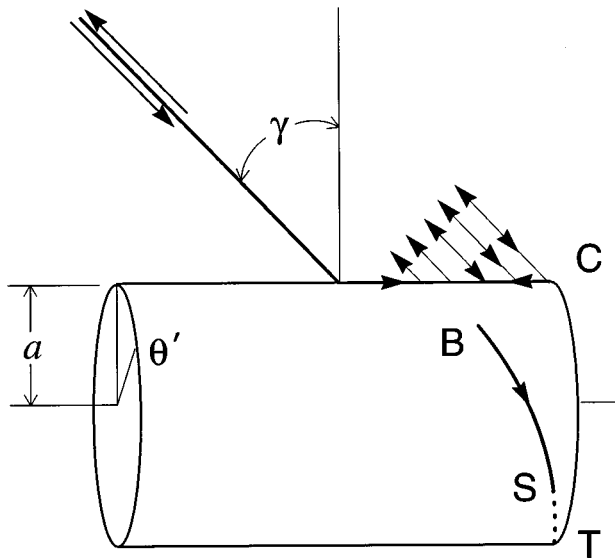
FIG. 1. Two leaky wave processes on a cylinder where a three-dimensional formulation is appropriate: (i) the tilt angle γ is at the leaky wave coupling angle θ_l so that the ray lies in the plane defined by the incident wave vector and the cylinder's axis; and (ii) regular helical ray excitation at a polar angle θ' meeting a stationary phase condition [Eq. (29)]. Backscattering enhancements can result for (i) from a reflection at C (Refs. 1 and 19) and for (ii) from a reflection at for example T (see Sec. V). Ray (i) is termed a meridional ray following standard optical terminology.

pling regions for plane and spherical waves incident on a flat surface according to the formulation given in Sec. I. Section III discusses the approximate generalization of the formulation given in Sec. I for flat surfaces to the case of curved scatterers. Section IV analyzes the effect of an idealized partial coating extending down the axis of a right circular cylinder as shown in Fig. 2. This coating will be called a *y* coating to distinguish it from the angular θ coating previously analyzed[18] (Fig. 2M) which was a two-dimensional problem unlike Fig. 2. While the analysis here has some similarities with a previous study of specular reflection by finite cylinders,[24] the process considered is different. Results are compared with other formulations for the uncoated cylinder case and in Sec. V the amplitudes and well-known coupling conditions for helical waves on a tilted circular cylinder are derived from the stationary phase condition for the integral introduced in Sec. I.

The formulation in Secs. I and III decomposes the complex wave-number dependence of the reflection coefficient in the vicinity of a leaky wave pole into a regular part and a term with a simple pole. The approach is similar to Bertoni and Tamir[25] in that the terms are identified with specular and leaky wave contributions to the outgoing wave. The pole is assumed to be such that the real component k_l of the leaky wave number is much larger than the imaginary component α. Appendix A comments on the residues needed for the evaluation of leaky wave amplitudes. Certain simplifications discussed there can introduce phase errors (in addition to those noted below) without significantly altering the estimated magnitude of the leaky wave. For many applications of interest, the observer is in the shadow zone of the specular
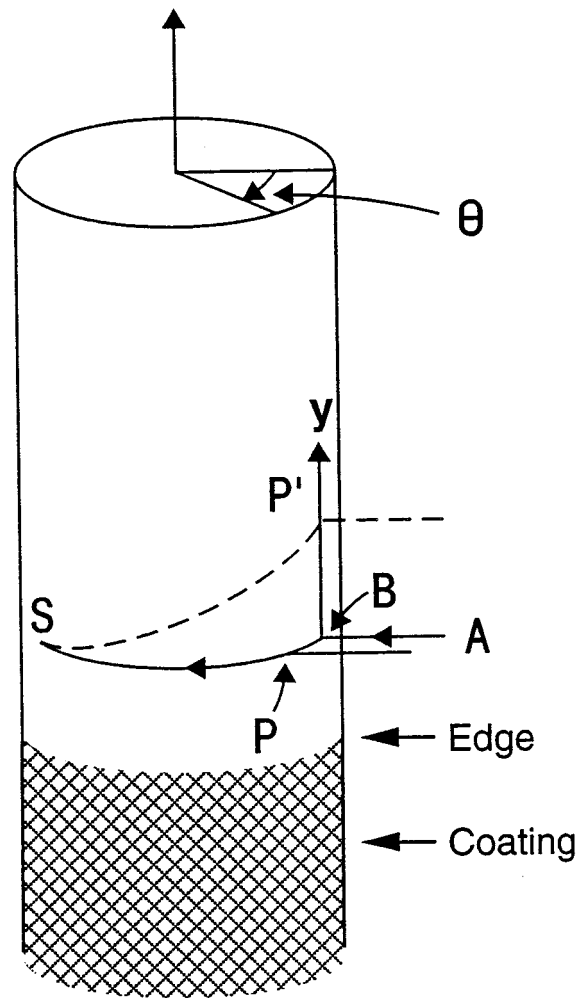
FIG. 2. Partially coated cylinder illuminated at normal incidence and considered in Sec. IV. The cylinder is taken to be infinitely long with the region $y \leqslant y_E$ coated to block the incident wave. The outgoing leaky wave amplitude at S is determined by integrating over paths that include points P' lying outside the plane of incidence. One effect of the coating is to introduce a helical raylike contribution excited with P' at the coating edge. The coating may be terminated with θ outside the illuminated region and its edge may lie above B so as to block the circumferential ray in the plane of incidence.

contribution or the specular contribution may be separated in time from the leaky wave. Since the leaky wave contribution is often primarily of interest for such applications, a description of the total outgoing amplitude uniform in the spatial variables is not considered. Examples of uniform descriptions of reflection from plane and/or cylindrical surfaces have been given by others.[4,26–28]

The leaky wave contribution to the detected signal depends not only on the outgoing amplitude at the surface and on the shape of the scatterer but also on the aperture and focal properties of the receiver.[29,30] Those system-specific aspects are not examined here though an example of ray propagation to a distant receiver is noted in Appendix B.

The method of approximating out-going leaky wave amplitudes given in Secs. I and III may introduce phase errors in the predicted amplitudes. The phase errors are wave kinematic in nature, being proportional to the wavefront and surface curvatures at launching and detachment points of the

leaky waves. These errors are evident from comparisons with other formulations[3–7] noted in Appendix B for the uncoated limit of the right circular cylinder. This class of phase error is $O(1/ka)$ for plane-wave incidence where a is the cylinder's radius and k is the wave number in water. Such phase errors can be unimportant for high-frequency sonar applications because the scattering may be dominated by a single ray process (as illustrated in Ref. 1 and Fig. 1) such that there is negligible interference between ray amplitudes. Thus certain phase differences need not be accurately approximated for estimating the target amplitude. Furthermore, evaluation of the radiation damping rates of leaky waves above the coincidence frequency[9–15] suggests that for many applications repeated reflections or circumnavigations of leaky waves across the long dimensions of a scatterer may be unimportant at high frequencies. For some decisions concerning target discrimination, even an order-of-magnitude estimate of the leaky wave amplitude may be adequate.

I. SUPERPOSITION OF TWO-DIMENSIONAL LEAKY WAVES ON FLAT SURFACES

In this section the leaky wave is excited on a flat isotropic surface at $z=0$ by an incident acoustic wave that has an amplitude $p_i(x,y,z)$ in the absence of the surface. A time factor of $\exp(-i\omega t)$ is used throughout. As reviewed in Appendix A, the amplitude reflection coefficient R is a function of the in-plane wave vector ζ such that for ζ in the vicinity of a complex leaky wave pole k_p, the following Laurent expansion is applicable:

$$R(\zeta \approx k_p) = R_0(\zeta) + [\text{Res}(R,k_p)/(\zeta - k_p)], \tag{1}$$

where R_0 is smooth near k_p and Res denotes the residue of R at the indicated simple pole. For $k_p = k_l + i\alpha$ with $k_l > 0$, then $\alpha > 0$ though R also has a symmetric pole at $-k_p$. The analysis given below allows the pressure contribution to the outgoing wave evaluated on the surface resulting from the leaky wave pole to be approximated at surface point S as

$$p_l(S) \approx \int_{\mathscr{D}} p_i(S')[\kappa H_0^{(1)}(k_p s)]d\mathscr{A}', \tag{2a}$$

$$\kappa = i \ \text{Res}(R,k_p)k_p/2, \tag{2b}$$

$$s = [(x-x')^2 + (y-y')^2]^{1/2}, \tag{2c}$$

where $H_0^{(1)}$ is the Hankel function having an argument proportional to the distance s between points S' at (x',y') and S at (x,y). The area differential $d\mathscr{A}'$ may be written $dx' \ dy'$ and the domain of integration $\mathscr{D}$ will be discussed in Sec. II. The contribution to the total outgoing amplitude p_{out} associated with (2) is only that of the isolated leaky wave, and other contributions such as lateral waves and specularly reflected waves are not considered here since they are taken to be distinct or unimportant for scattering processes of interest (e.g., Fig. 2). The arguments leading to (2) are perhaps more similar to methods used in the Fourier analysis of optical and acoustical imaging systems[29,30] and in ultrasonic nondestructive testing[26,27] than in common formulations of scattering theory. The underlying object is not restricted to an elastic half space since it could also be a plate,

a layered structure, or a fluid waveguide.[31] It may dissipate or thermalize energy.

The spatial spectrum of the incident wave evaluated on the surface is

$$P_i(k_x,k_y) = \frac{1}{2\pi} \int \int_{-\infty}^{\infty} p_i(x,y,0)e^{-ik_x x - ik_y y} \ dx \ dy, \tag{3}$$

and the total outgoing wave is

$$p_{\text{out}}(x,y,z) = \frac{1}{2\pi} \int \int_{-\infty}^{\infty} [P_i(k_x,k_y)R(\zeta)]$$
$$\times e^{i(k_x x + k_y y + k_z z)} \ dk_x \ dk_y, \tag{4}$$

$$\zeta = (k_x^2 + k_y^2)^{1/2}, \quad k_z = (k^2 - k_x^2 - k_y^2)^{1/2}, \tag{5}$$

where $k = \omega/c$ and c is the phase velocity in the surrounding fluid. Equation (2) is obtained by evaluating the leaky wave contribution to the outgoing wave at $z=0$ for the case of an ''input'' $[p_i(x,y,z=0) \equiv p_i(x,y)]$ given by a two-dimensional delta function: $\delta(x,y) = \delta(x)\delta(y)$. The resulting spectra from (3) is $P_i = 1/2\pi$. Letting $r = (x^2+y^2)^{1/2}$ and $k_r = \zeta$, Eq. (4) may then be simplified at $z=0$ to the following Hankel transform of zero order:

$$p_{\text{out}\delta}(x,y,0) = \frac{1}{2\pi} \int_0^{\infty} R(k_r)J_0(k_r r)k_r dk_r, \tag{6}$$

by introducing polar coordinates and integrating over the angle variable (see, e.g., Ref. 31). This may be rewritten by using the symmetry of the general form of $R(-\zeta) = R(\zeta)$ and by taking $J_0(k_r r) = [H_0^{(1)}(k_r r) + H_0^{(2)}(k_r r)]/2$. The symmetry properties of the integrand (as reviewed, e.g., by Frisk[31]) are such that

$$p_{\text{out}\delta}(r) = \frac{1}{4\pi} \int_{-\infty}^{\infty} R(k_r)H_0^{(1)}(k_r r)k_r \ dk_r, \tag{7}$$

with the path of integration slightly displaced into the second and fourth quadrants in order that the radiation condition be satisfied in the evaluation of (4). To restrict attention to the pole contribution, Eq. (1) is used and the path is deformed to a contour $\mathscr{C}$ which accounts for branch lines. The resulting leaky wave contribution to the spatial response is

$$p_{l\delta}(r) = \frac{1}{4\pi} \int_{\mathscr{C}} \frac{\text{Res}(R_1,k_p)}{(k_r - k_p)} k_r H_0^{(1)}(k_r r)dk_r \tag{8a}$$

$$= \kappa H_0^{(1)}(k_p r), \tag{8b}$$

where (8b) follows since $\mathscr{C}$ is taken to enclose the indicated pole. Since (8b) gives the leaky wave responses from a δ function ''input,'' (2a) follows from superposition where $\mathscr{D}$ is selected so as to approximate the radiating components in the direction of interest as explained in Sec. II. A heuristic derivation of κ is to enforce the recovery of a special case previously considered[18,29] and shown in Sec. II.

While not emphasized here, the approximation (2) may have applications to scanning acoustic microscopy[29,30] and ultrasonic nondestructive testing beyond the examples in Sec. II.

II. PLANE SURFACE EXAMPLES: PARABOLIC AND OVAL FRESNEL PATCHES

While (8b) corresponds to an isotropic damped cylindrical wave diverging from each point of excitation, the domain $\mathscr{D}$ of the convolution in (2) is bounded to account for the assumed directionality of the outgoing wave of interest. This may be demonstrated by considering the special case where the incident wave does not depend on y as in previous convolution formulations.[18,29] Taking $y=0$, the y' integration in (2) becomes

$$\int_{-\infty}^{\infty} H_0^{(1)}(k_p(y'^2+(x-x')^2)^{1/2})dy' = \frac{2}{k_p}e^{ik_p|x-x'|}, \quad (9)$$

where it is assumed that $\text{Im}(k_p)=\alpha>0$ as in the situation under consideration; (9) follows from the Fourier inverse of an integral representation of $H_0^{(1)}$ evaluated at zero spatial frequency (see, e.g., Ref. 32). Consider the situation where the leaky wave component of interest is propagating in the direction of increasing x. The previous results[18,29] are recovered if $\mathscr{D}$ is restricted to the region with $x'\leq x$ so that (2) gives

$$p_l(x) \approx i\, \text{Res}(R,k_p)e^{ik_px}\int_{-\infty}^{x}p_i(x')e^{-ik_px'}\,dx'. \quad (10)$$

From Eq. (A5), the coefficient reduces to the corresponding results in Refs. 18 and 29 for the case of a lossless elastic surface when the residue is approximated to leading order in α/k_l. The restriction of the domain follows from Cauchy's theorem and the required choice of contour.[18,29] The restriction is consistent with Stamnes' analysis[33] of the insignificance of backward propagating leaky wave contributions for bounded beam illumination. The pathological case of leaky waves having opposite group and phase velocities observed for example in Ref. 13 is not considered here.

For three-dimensional incident waves, the above result for $\mathscr{D}$ is generalized without proof: To approximate the leaky wave component propagating in the $+x$ direction, (2) is evaluated by integrating x' from $-\infty$ to x and y' from $-\infty$ to ∞. [Restricting $\mathscr{D}$ in this way is equivalent to convolving $p_i(x,y)$ over an infinite $\mathscr{D}$ with a directionally limited response function.] This restriction is consistent with the analysis of Fresnel coupling regions as will now be illustrated. Rays and Fresnel zones are found by replacing $H_0^{(1)}$ by its large argument asymptotic form giving

$$p_l(S) \approx \kappa\left(\frac{2}{\pi}\right)^{1/2}e^{-i\pi/4}\int_{\mathscr{D}}p_i(S')(k_ps)^{-1/2}e^{ik_ps}\,d\mathscr{A}', \quad (11)$$

where the approximation breaks down in the region of $\mathscr{D}$ where the surface path length s is not large. To identify the surface regions contributing significantly to p_l, it is helpful to first identify any Fermat paths corresponding to S' for which the phase of $p_i(S')\exp(ik_ls)$ is stationary for a given output surface point S. Let such points S' be denoted by B. The path through B to S represents a *ray* to S. The locus of surface points S' for which the *phase defect*

$$\text{Arg}(p_i(S')e^{ik_ls})-\text{Arg}(p_i(B)e^{ik_l\sigma})\equiv\Delta\varphi=\pm\pi \quad (12)$$
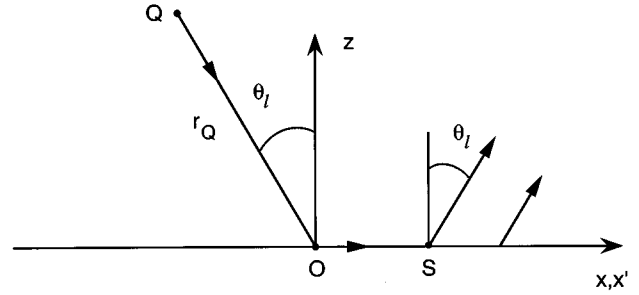
FIG. 3. Leaky wave excitation geometry considered in Sec. II. For plane-wave incidence the angle of incidence is taken to be $\theta_l=\sin^{-1}(k_l/k)$. For a source at Q, the origin O is chosen so that QO is at θ_l from the normal for the purpose of calculating the right-directed leaky wave amplitude radiated at S.

defines the boundary of the Fresnel coupling patch for each stationary point under consideration. Here σ is s evaluated at $S'=B$ and $\Delta\varphi$ vanishes as S' approaches B. Adjacent surface points more distant from B than this Fresnel locus interfere destructively with the contribution from the Fermat path in the usual case where $|\Delta\varphi|$ increases monotonically with the distance between S' and B. Some regions outside the Fresnel patch are included in $\mathscr{D}$ to avoid overestimating $|p_l|$. While (12) is the simplest generalization of the previous analysis of Fresnel patches for spheres and cylinders,[17,18] there can be situations where the extent of the significant coupling region is strongly affected by the damping α as illustrated below for plane-wave incidence on a flat surface.

Consider the situation illustrated in Fig. 3 for the case of plane-wave incidence at the real-valued leaky wave coupling angle $\theta_l=\sin^{-1}(k_l/k)$ such that $p_i=p_o\exp(ik_lx')$. It is convenient to introduce normalized coordinates $\bar{x}'=k_lx'$, $\bar{y}'=k_ly'$, $\bar{x}=k_lx$, and $\bar{y}=k_ly$ and to take $y=0$ at S. The Fermat path is degenerate and includes all points on the x' axis to the left of x. The Fresnel boundary from (12) is a parabola which opens to the left:

$$(\bar{x}-\bar{x}')+\pi/2=\bar{y}'^2/(2\pi), \quad (13)$$

shown in Fig. 4. Coupling paths along this parabolic locus are delayed by π radians relative to the set of Fermat paths. Figure 4 also shows a circular arc of radius $\bar{r}=k_l/\alpha$ centered on S to show the locus of coupling paths beyond which surface propagation to the point S of interest attenuates the amplitude in excess of e^{-1}. The right-directed outgoing leaky wave amplitude at S is relatively weakly influenced by coupling paths outside of these parabolic and circular curves. A partial coating of the interface which blocks some of the illumination outside that patch is anticipated to have a weaker effect than if a corresponding area were blocked within the patch. (It is assumed here that the coating only weakly affects the local value of α so that only changes in the aperture are under consideration.) If the *paraxial approximation* $s\approx|x-x'|+(y'^2/2|x-x'|)$ is used in (12), the Fresnel locus is parabolic with the vertex shifted to x giving (13) without the $\pi/2$ term. Restricting $\mathscr{D}$ as in (10) omits only a small portion of Fresnel patch from (12) near the vertex, and none of the Fresnel patch is omitted if the paraxial approximation for s is used.
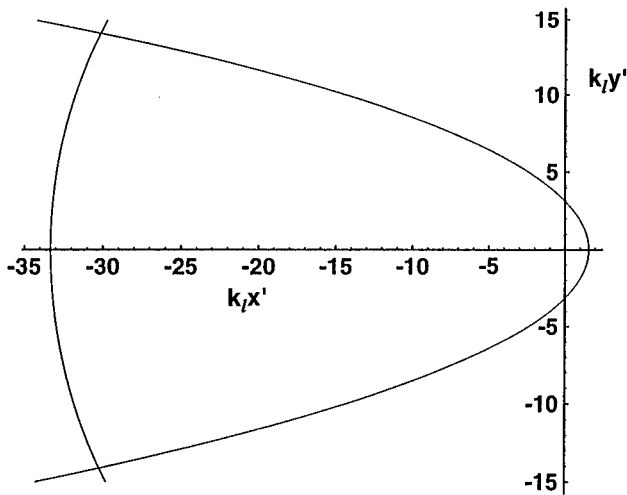
FIG. 4. The parabola shows the Fresnel patch boundary for a surface point S at the origin given a plane-wave incident from the left with an angle of incidence equal to θ_l as in Fig. 3. The circular arc is the boundary for e^{-1} attenuation in propagation to S for the case where $\alpha/k_l = 0.03$ (corresponding approximately to leaky Rayleigh waves on an aluminum–water interface). For smaller values of α/k_l (i.e., a steel–water interface) the radius of the arc is larger in proportion to k_l/α. Blocking the illumination outside the region bounded by these curves has a relatively weak effect on the amplitude at S (see text). The phase defect increases rapidly for $x' > 0$.

Evaluation of (10) with $p_i(x') = p_o \exp(ik_l x')$, corresponding to the case considered in Fig. 4, gives

$$p_l(x) \approx i p_i(x) \mathrm{Res}(R, \xi_p)/\alpha \approx -2 p_i(x) R_0(k_l), \qquad (14)$$

where the right-most expression uses Eq. (A5) and omits an $O(\alpha/k_l)$ correction to the residue for an elastic surface fluid-loaded on only one side. As explained in Appendix A, (A5) can be a useful approximation for leaky Rayleigh waves or for leaky Lamb waves on a plate fluid-loaded on only one side. Approximating the specular reflection contribution as $p_{\text{spec}} \approx p_i(x) R_0(k_l)$ yields $p_l(x) \approx -2 p_{\text{spec}}$ in agreement with an analysis of a wide beam incident at the Rayleigh angle.[34] [The total reflected field becomes $p_{\text{out}} \approx -p_i(x)$; this is expected since $R(k_l) \approx -1$.] Evaluation of (10) for the special case of an exponentially weighted beam recovers results from Ref. 25 (Sec. 3.1) when (A5) is also used.[18]

Consider now the Fresnel patch for a point source Q (see Fig. 3). The origin O may be chosen such that the line QO makes an angle $\theta_l = \sin^{-1}(k_l/k)$ with respect to the surface normal. The length QO, denoted by r_Q, is such that $kr_Q \gg 1$. The source and surface point of interest S lie in the plane $y = 0$, as does the Fermat path which is at $x' = 0$ for all $x \geq 0$. Equation (12) gives the following Fresnel locus:

$$[(\bar{x}^2 - 2\bar{x}x' + \bar{s}'^2)^{1/2} - \bar{x}]$$
$$+ [(\bar{r}_Q^2 + \bar{s}'^2 - 2\bar{r}_Q \bar{x}' \sin \theta_l)^{1/2} - r_Q]/\sin \theta_l = \pi, \qquad (15)$$

where $s' = (\bar{x}'^2 + \bar{y}'^2)^{1/2}$ and $\bar{r}_Q = k_l r_Q$. Figure 5 shows several Fresnel patches for $\bar{x} = 15$, which is distant from the transition region at $\bar{x} = 0$. The patches are oval shaped, approaching the parabolic locus (13) from the left as r_Q becomes large. For surface points outside the patch, the phase
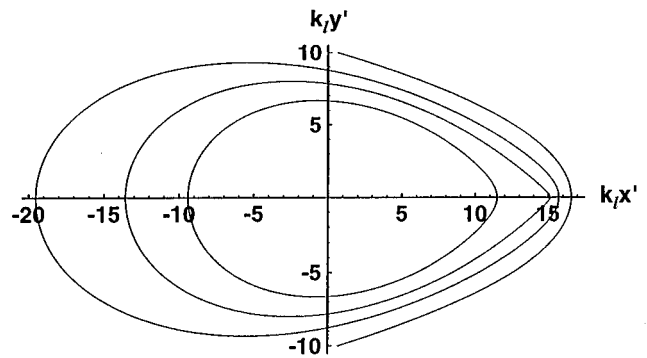
FIG. 5. Oval Fresnel patches with normalized source distances $k_l r_Q$ of 25, 50, and 100 (beginning with the smallest oval) for the geometry shown in Fig. 3 with $\theta_l = 30°$. The surface point S is at $k_l x = 15$, and the stationary phase point is at the origin. For sufficiently small r_Q, the patch is entirely to the left of S. When $k_l r_Q$ becomes very large, the right side of the oval approaches the parabola shown given by Eq. (13).

of the integrand in (11) varies rapidly, giving a weak contribution to the integral. As $\bar{r}_Q$ decreases, the right vertex of the oval has $\bar{x}' < \bar{x}$ which supports the omission of the region having $x' > x$ from $\mathscr{D}$ in (2). When r_Q is sufficiently large, the attenuation α also affects the left-most boundary of the region contributing to (2) as may be seen by superposing on Fig. 5 a circle of radius k_l/α centered on $(\bar{x}, \bar{y}) = (15, 0)$. The patch area decreases with decreasing r_Q which results in a corresponding reduction in $|p_l|$ provided the source strength is decreased so as to keep $|p_i(0)|$ constant.

III. GENERALIZATION OF EQ. (2) TO CURVED SURFACES

For fluid-loaded curved surfaces, k_p is taken to be given locally from continuum mechanics. For nondissipative surfaces, $\mathrm{Res}(R, k_p)$ in κ is approximated by Eq. (A5) or (A7), as illustrated by the example in Appendix B. Attention is restricted to surfaces for which the dependence of k_l and α on the changes in curvature and on the direction of propagation have a negligible effect on the amplitudes of interest in comparison to the dephasing effects discussed below. The existence of high-frequency regimes where that assumption will be useful for estimating local leaky wave amplitudes is supported by calculations of leaky wave properties for canonical thick shells (with some comparisons to normalized plate theory) and other solids (see, e.g., Refs. 10–13, 16, and references cited therein). Related simplifications have been implemented in other formalisms even below the coincidence frequency (e.g., Refs. 6–8). The principle effect of surface curvature on the excitation of the leaky wave becomes the modification of the spatial dependence of the incident amplitude p_i in Eq. (2). In the evaluation of (2a), the geodesic surface path length along the flat surface (2c) is replaced by the arc length for the geodesic between S and S' of the curved surfaces. The integration domain $\mathscr{D}$ used in the approximation of p_l may be truncated as discussed in Sec. II.

The examples given in Secs. IV and V are restricted to situations where there are isolated leaky wave Fermat paths to S. There are potentially important scatterers not having this property (see, e.g., M Sec. IV).
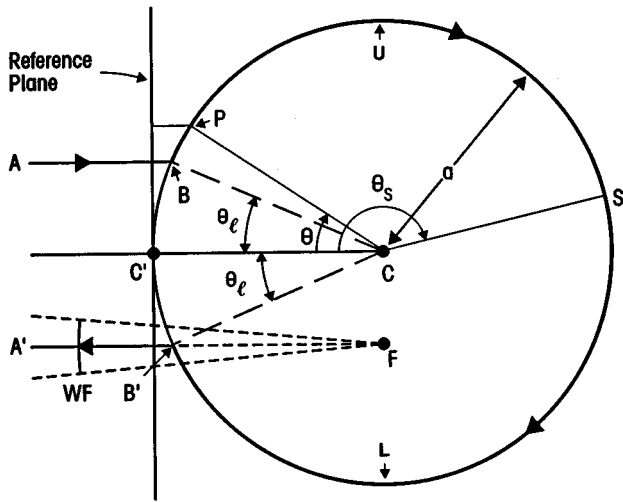
FIG. 6. Ray diagram for a circular cylinder in the plane of incidence showing only the clockwise propagation. Points U and L denote the upper and lower boundaries of the illumination. Moving S around to B' leads to a backscattered ray and the diverging ray tube used in the calculation of Eq. (B4) is shown. The diagram also applies to the projection of helical ray paths (Figs. 1 and 8) provided the angles of B and B' are replaced by θ' from Eq. (29a) and projection beyond $\theta = \pi$ is taken to be the end-reflected ray discussed in Sec. VI.

IV. *y*-COATED CYLINDERS AT NORMAL INCIDENCE: EDGE-EXCITED HELICAL WAVES

A. Formulation

Consider the situation depicted in Fig. 2 with plane-wave normal incidence. The surface may be unwrapped into a plane as discussed, e.g., by Pierce and Kil[35] where in the notation of Fig. 2 the coordinates are (x,y) with $x=\theta a$. The line $y=0$ corresponds to the intersection of the surface with the *plane of incidence* in Fig. 2. The formulation from Sec. III allows for the paths outside the plane of incidence in the approximation of the leaky wave amplitude at S. Figure 6 shows rays and paths in the plane of incidence. For brevity, attention is restricted to leaky waves propagating in the direction of increasing θ with an aperture for excitation restricted to $\theta = -\pi/2$ to $\pi/2$ corresponding to the region illuminated by the incident wave and using the Kirchhoff approximation. The surface coordinate $x_S = a\theta_S$ at S is such that $\theta_s > \pi/2$ but may otherwise be arbitrarily large. The effect of repeated circumnavigations and the modulo $(2\pi a)$ illumination apertures may be accounted for by superposition if that is important for the application of interest. For an infinite cylinder with a coating edge at y_E, (2) gives

$$p_l(S) \approx \kappa \int_{x_L}^{x_U} p_i(x')dx' \int_{y_E}^{\infty} H_0^{(1)}(k_p s)dy', \qquad (16a)$$

$$s = [(x-x')^2 + y'^2]^{1/2}, \qquad (16b)$$

$$p_i = p_B \exp[ika(\cos\theta_l - \cos x'/a)], \qquad (16c)$$

where $y=0$ at S and the effects of the idealized coating considered are limited to the restriction of the illuminating aperture and $x_{U,L} = \pm a\pi/2$. The incident amplitude is p_B at the Fermat coupling point B at $\theta_l = \sin^{-1}(k_l/k)$. The following parameters occur in the evaluation of (16) for an un-

coated cylinder given in Appendix B which are also used in the approximations discussed below:

$$\Delta\theta_f = [2\pi/(ka\cos\theta_l)]^{1/2}, \qquad (17a)$$

$$\beta_f = a\alpha\Delta\theta_f, \qquad (17b)$$

$$b = \beta_f/(2\pi^{1/2}), \qquad (17c)$$

where $\Delta\theta_f$ is the *Fresnel angular half-width* of the coupling region in radians[17] and β_f is the damping rate relative to the Fresnel half-width.[18] It is assumed below that ka is large enough for $|x_{L,U} - a\theta_l| \gg \Delta\theta_f$ with $b \ll 1$. The discussion in Appendix B is more general and may be read first.

B. Paraxial approximation and the Fresnel area

In (16), the approximation $H_0^{(1)}(k_p s) \approx (2/\pi k_p s)^{1/2} \times \exp(ik_p s - i\pi/4)$ may be used and the dominant contributions to the integral are from regions where the paraxial approximation $s \approx (x-x')(1 + y'^2/2(x-x')^2)$ may be used in the phase factor (see the discussion at the end of Sec. IV). The $s^{-1/2}$ spreading factor may be approximated as $(x-x')^{-1/2}$ which reduces (16) to

$$p_l \approx \kappa(2/\pi k_p)^{1/2}e^{-i\pi/4}I_{xy}, \qquad (18a)$$

$$I_{xy} = \int_{x_L}^{x_U} p_i(x')(x-x')^{-1/2}e^{ik_p(x-x')}I_y(x')dx', \qquad (18b)$$

$$I_y(x') = \int_{y_E}^{\infty} e^{ik_p y'^2/2(x-x')}dy', \qquad (18c)$$

where p_i is given by (16c). For the uncoated cylinder y_E goes to $-\infty$ and the domain of integration in I_{xy} contains the two-dimensional stationary phase point at $(x',y') = (a\theta_l,0)$. In the paraxial approximation, the boundary of the Fresnel coupling patch is an ellipse with principal axes in the x and y directions. The half-width in x' is $a\Delta\theta_f$ and the axial Fresnel half-width is[17]

$$\Delta y_f = [2\pi a(\theta_S - \theta_l)/k_l]^{1/2}, \qquad \theta_S \gg \theta_l, \qquad (19)$$

which depends on the angle θ_S at the surface point of interest. In the high-ka limit, where the integration domain encloses several Fresnel zones, for weak damping (see below) I_{xy} may be approximated by the two-dimensional stationary phase approximation (SPA) which gives[20,22]

$$I_{xy} \approx p_B(a\theta_S - a\theta_l)^{-1/2}e^{i(\varphi_{BS} + \pi/2)}A_f/\pi, \qquad (20a)$$

$$A_f = \pi a\Delta\theta_f\Delta y_f, \qquad (20b)$$

where A_f is the area of the first elliptical Fresnel zone (in the quadratic phase approximation) and $\varphi_{BS} = (k_l + i\alpha)a(\theta_S - \theta_l)$. Equations (18)–(20) may be combined to give

$$p_l \approx i \, \text{Res}(R,k_p)ap_Be^{i\varphi_{BS}}(k_p/k_l)^{1/2}\Delta\theta_fe^{i\pi/4}, \qquad (21)$$

which corresponds to (B1a) with the appropriate limiting form for I [from below Eq. (B2)] with modifications to the magnitude and phase that are small for weak damping.
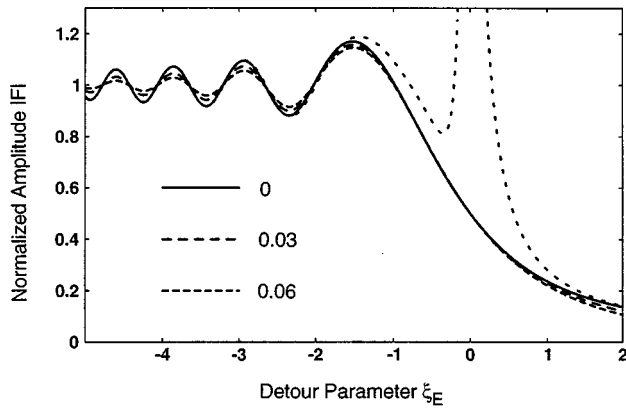
FIG. 7. The magnitude of the Fresnel integral function $F(\xi_E)$ is shown as the solid curve. The dashed curves with the indicated values of α/k_l are given by Eq. (24). They describe the approximate magnitude of the correction factor to account for the coating shown in Fig. 2. The detour parameter ξ_E depends on both the edge location relative to the surface point S (at $y = 0$) and on the Fresnel half-width Δy_f which depends on the location of S. Unlike the case of $\alpha/k_l = 0.03$ shown for leaky Rayleigh waves on an aluminum–water interface, the deviation from the solid curve is small for a steel–water interface which has $\alpha/k_l \approx 0.01$. The asymptotic approximation of $|F(\xi_E)|$ is given by the dotted curve for the case $\alpha = 0$ (see text).

C. y-coated cylinder

Under conditions to be discussed subsequently, the approximations leading to (18) apply. Evaluation of (18c) gives[36]

$$I_y(x') = e^{i\pi/4}[2\pi(x-x')/k_p]^{1/2}\{\tfrac{1}{2}\,\mathrm{erfc}[(k_p/2(x-x'))^{1/2}y_E e^{-i\pi/4}]\}, \quad (22)$$

where the phase of the argument of erfc is such that $I_y(x')$ is a slowly varying function near $x' = x_l \equiv a\theta_l$. It is assumed that $x_l - x_L$ and $x_U - x_l$ are sufficiently large that (18b) may be approximated by stationary phase giving

$$I_{xy} \approx ap_B e^{i\varphi_{BS}}\Delta\theta_f e^{i\pi/4}(x-x_l)^{-1/2}I_y(x_l), \quad (23)$$

so that (18a) becomes

$$p_l \approx i\,\mathrm{Res}(R,k_p)ap_B e^{i\varphi_{BS}}(k_p/k_l)^{1/2}\Delta\theta_f e^{i\pi/4}F_E, \quad (24a)$$

$$\xi_E = \pi^{1/2}y_E/\Delta y_f, \quad (24b)$$

$$F_E = F(\xi_E(k_p/k_l)^{1/2}) = \tfrac{1}{2}\,\mathrm{erfc}[e^{-i\pi/4}\xi_E(k_p/k_l)^{1/2}], \quad (24c)$$

$$F(\tau') = \frac{e^{-i\pi/4}}{\pi^{1/2}}\int_{\tau'}^{\infty} e^{-i\tau^2}\,d\tau, \quad (24d)$$

where F denotes *a Fresnel integral of complex argument* and $(k_p/k_l)^{1/2} \approx 1 + i(\alpha/2k_l)$. Figure 7 shows the transition in $|F_E|$ that occurs as the *detour parameter*[22] ξ_E for the coating edge varies. (The terminology parallels that of M Sec. III for the θ-coated cylinder though the complex argument of F differs in form.) The leading asymptotic terms of $F(\tau)$ as $|\tau| \to \infty$ are $F(\tau) \sim H(\tau) + F_1(\tau)$ with[37]

$$F_1(\tau) = (2\pi^{1/2}\tau)^{-1}\exp(i\tau^2 + i\pi/4),$$

$$H(\tau) = \begin{cases} 0, & \text{if } (-\pi/4) < \arg \tau < (3\pi/4), \\ \tfrac{1}{2}, & \text{if } \arg \tau = (-\pi/4) \text{ or } (3\pi/4), \\ 1, & \text{if } (3\pi/4) < \arg \tau < (7\pi/4). \end{cases} \quad (25)$$

These yield the following ray interpretation when it is noted that $i\tau^2 = i\xi_E^2 k_p/k_l$ in F_1 is the phase and attenuation defect of a ray which would not be present without the coating. That ray is excited by the acoustic field *at the edge of the coating* at a point P' with $\theta = \theta_l$ and $y = y_E$ (compare Fig. 2). The ray travels a *geodesic helical path* to the surface point S and the phase and attenuation defect are relative to a Fermat path through B. The edge-ray amplitude is asymptotically described by $F_1(\xi_E(k_p/k_l)^{1/2})$, which may be used to formulate a complex *diffraction coefficient*.[22,37] The term $H(\tau)$ describes the presence or absence of the Fermat leaky ray amplitude according to whether the Fermat path through B is illuminated (so that $\xi_E < 0$ and $H = 1$) or blocked (so that $\xi_E > 0$ and $H = 0$). When $\xi_E \ll -1$, the structure in F_E evident in Fig. 7 corresponds to the interference between the ordinary and edge rays. The relative edge-ray amplitude is reduced by increasing α/k_l because the helical path is longer than the reference path BS. Therefore the larger the attenuation the more the interference structure is suppressed.

Figure 7 shows how both the total amplitude and the approximate edge-ray contribution vary with $y_E/\Delta y_f$. The coating is tangent to the elliptical boundary when $|\xi_E| = \pi^{1/2}$ and it is necessary to illuminate a surface region somewhat wider than the Fresnel zone to recover the uncoated leaky wave amplitude. Constructive interference between the direct surface ray and the edge ray contributions can produce a significant enhancement for the partially coated case relative to the uncoated cylinder. For $|\xi_E| < \pi^{1/2}$, the dotted and solid curves in Fig. 7 separate from each other because the helical edge-ray contribution is no longer distinct from the direct ray contribution.[22]

D. Assumptions and approximations

The physical situation considered is that the dominant effect of the coating is to limit the illuminated aperture for the excitation of leaky waves on the cylinder. Real coatings may have additional effects that may need to be included. For example, leaky ways may be weakly reflected from the coating edge or, for the case where $\xi_E > 0$, the attenuation of leaky waves which must travel under a coated region to reach point S, may cause k_p to be altered by the coating. For situations where the fluid loading is weak so that α/k_l is $\ll 1$, those affects are probably small.

Some discussion of the paraxial approximation is appropriate. It is necessary for the resulting phase errors[22] to be insignificant in the dominant regions of integration in (16). That needs to be the case over at least the Fresnel width Δy_f which leads to the condition $a^3(\theta_S - \theta_l)^3 \gg (k_l/8)(\Delta y_f)^4$, which becomes $k_l a(\theta_S - \theta_l) \gg \pi^2/2$. For the interference structure in Fig. 7 to be accurately described beyond the Fresnel width, it is necessary that $a^3(\theta_S - \theta_l)^3 \gg (k_l/8)y_E^4$; however, the magnitude of the edge-ray contribution falls off rapidly for increasing $|y_E|$. Finally, comparison with the analysis for the uncoated cylinder, Eqs. (B1) and (B2),

1634   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Philip L. Marston: Scattering from cylinders   1634
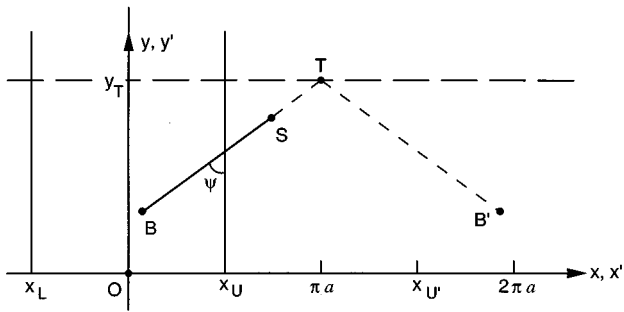
FIG. 8. Unwrapped plane surface for the description of regular helical ray excitation on the tilted cylinder shown in Fig. 1. The y axis is parallel to the cylinder's axis as in Fig. 2. The Fermat point B and amplitude for a given surface point S are analyzed in Sec. VI. The dashed ray extended beyond S is shown to reflect from a truncation at T to reach B', the image of B reflected about the line $x = \pi a$. See also Fig. 6.

shows that (21) and (24) lack a factor of $\exp(ib^2)$; from (17), b is almost always small because the damping over the angular Fresnel width $\Delta \theta_f$ is usually small.

V. HELICAL WAVE EXCITATION ON INFINITE AND TRUNCATED TILTED CYLINDERS

It is well known that matching of the local surface wave-vector component with the tangential component of the incident wave vector k_i on a tilted cylinder can excite helical waves.[1,3–6,38] The usual matching conditions for isotropic waves, Eqs. (29a) and (29b) below, are derived here as a consequence of the form of Eq. (2) for the parameters shown in Fig. 1. Figure 8 shows the unwrapped Cartesian coordinates used in the analysis, where the illuminated domain of interest is $x_L < x' < x_U$, $x_{U,L} = \pm \pi a/2$. The origin O lies in the plane containing k_i and the cylinder's axis so that in the illuminated strip the incident wave is

$$p_i(x',y') = p_O e^{i\varphi_i}, \tag{26a}$$

$$\varphi_i = ka(1 - \cos(x'/a))\cos \gamma + ky' \sin \gamma, \tag{26b}$$

where γ is the tilt angle (Fig. 1). Approximating $H_0^{(1)}$ in (2) for large $|k_p s|$ as in (11) gives an integrand having the phase

$$\varphi_t(x',y';x,y) = \varphi_i(x',y') + k_l[(x-x')^2 + (y-y')^2]^{1/2}, \tag{27}$$

where x and y denote the surface point S of interest. Attention is restricted to the case $x > x_U$, $y \geqslant 0$, so that S lies to the right of the illuminated strip. The Fermat path corresponds to the stationary phase path of all the paths summed by the integral in (11). The stationary conditions $\partial \varphi_t / \partial x' = 0$ and $\partial \varphi_t / \partial y' = 0$, respectively, give

$$(k_l/s)(x-x') = k \cos \gamma \sin \theta', \tag{28a}$$

$$(k_l/s)(y-y') = k \sin \gamma, \tag{28b}$$

where $\theta' = x'/a$ is the polar angle at the stationary point. Let $\psi = \tan^{-1}((x-x')/(y-y'))$ denote the pitch angle of the helical Fermat ray as shown projected on the plane in Fig. 8. Elimination of unknowns from (28) gives the desired coupling conditions

$$\sin(\theta') = (\cos \gamma)^{-1}(\sin^2 \theta_l - \sin^2 \gamma)^{1/2}, \tag{29a}$$

$$\cos \psi = \sin \gamma / \sin \theta_l, \tag{29b}$$

and the requirements that $\gamma \leqslant \theta_l = \sin^{-1}(k_l/k)$ and $\theta' \leqslant \theta_l$.

Following the notation of (12), let B denote the point (x',y') where the ray is launched toward S and let σ denote the ray length from B to S. To approximate the area of the elliptical Fresnel coupling patch, it is convenient to paraxially approximate the general path lengths relative to an axis along the Fermat path and to approximate the phase shift to second order in the deviation of (x',y') from B. The resulting Fresnel area and outgoing leaky wave amplitude at the surface are

$$A_f = (2\pi^2/\sin \psi)(\sigma a/k_l k \cos \theta' \cos \gamma)^{1/2}, \tag{30a}$$

$$p_l \approx \kappa(2/\pi k_p)^{1/2} p_B \sigma^{-1/2} e^{i\varphi_{BS} + i\pi/4} A_f / \pi, \tag{30b}$$

where $\varphi_{BS} = (k_l + i\alpha)\sigma$ and p_B is the incident wave amplitude phase referenced to the Fermat launching point B. The factors in (30b) are derived as in (18a) and (20a) and it is assumed that ka is so large that the fully illuminated Fresnel patch is not close to the edges of the illuminated strip. Though A_f increases in proportion to $\sigma^{1/2}$, that factor cancels in p_l so that the variation in $|p_l|$ with helical ray length σ is simply $\exp(-\alpha\sigma)$. Comparing (30) and (20b) shows that for fixed σ, tilting the cylinder alters A_f by a factor $(\cos \theta_l / \cos \theta' \cos \gamma)^{1/2} / \sin \psi = [1 - (\sin \gamma / \sin \theta_l)^2]^{-1/2} \geqslant 1$. The final result (30) is also given directly by the two-dimensional stationary phase approximation of (11). As in the case of the broadside examples given in Sec. IV, a weak dependence of phase on the attenuation rate is not included in the derivation.

Consider the magnitude of the far-field amplitude $|p_{lff}|$ resulting from a given helical wave surface amplitude p_l on an infinite cylinder. Construction of the relevant ray tube shows that $|p_{lff}| \approx (a \cos \theta_l / r)^{1/2} |p_l|$ for the contribution by the unidirectional helical wave where r denotes the radial distance of the far-field observation point from the cylinder's axis. From Eq. (29) it may be shown that $\cos \theta' \cos \gamma = \cos \theta_l$. When (A5) is used for κ in (30) and it is assumed that $\alpha \ll k_l$, the far-field amplitude becomes

$$|p_{lff}| \approx 2\alpha a |p_B| e^{-\alpha\sigma} (2\pi/kr)^{1/2} / \sin \psi. \tag{31}$$

This result reduces to the corresponding result in Eq. (91) of Norris and Rebinsky[6] for the special case of compressional membrane helical waves on a tilted thin cylindrical shell. The dependence on γ of the factor $1/\sin \psi$ is evident from $1/\sin \psi = [1 - (\sin \gamma / \sin \theta_l)^2]^{-1/2}$ which follows from Eq. (29).

Now suppose the cylinder is truncated along y_T in the plane of Fig. 8. The leaky wave may be reflected from this truncation as indicated by the dashed ray. Suppose the specular point is at $(\pi a, y_T)$ so that a point B' on the reflected ray is the image of B reflected about the line $x = \pi a$. The radiation from B' is found to be directed back towards the source which causes a backscattering enhancement as discussed, e.g., by Bao.[38] To approximate p_l at B' from (30), a helical wave reflection coefficient must be included and σ becomes the length of ray BTB'. The far-field amplitude will depend on the superposition of less-symmetric end-reflected helical rays and on the cylinder's length.

VI. DISCUSSION AND SCALING CONSIDERATIONS

The present formulation overcomes restrictions on target shape implicit in the early formulation by the present author[9,15,18] and examples in Secs. II, IV, and Appendix B recover previous amplitude predictions in the appropriate limiting cases. The analysis of the y-coated cylinder (Figs. 2 and 7) shows that helical waves can be launched at the edge of a coating at broadside incidence. The helical waves launched by this diffractive process are diminished in amplitude by the τ^{-1} dependence in the asymptotic approximation, Eq. (25). For the more common helical wave process examined in Sec. V, the leaky wave amplitude spreading factor $\sigma^{-1/2}$ is canceled by a corresponding increase in the Fresnel area. Circumferential waves for broadside illumination of the uncoated cylinder also exhibit this cancellation in Eq. (21). The Fresnel area is decreased by a more rapid dephasing along defective paths, which is the usual consequence of increasing the curvature of the surface or of the incident wavefront.

In high-frequency sonar systems it is possible to have both $\alpha L \gtrsim 1$ and $\alpha/k_l \ll 1$ where L denotes a dimension of the scatterer proportional to the length of the important surface rays. This is because typically $k_l L \gg 1$ and typically α increases in proportion to k_l. For real targets there will be dissipative contributions to α [which are permitted in (1) and (2) but not in (A3)–(A5)] in addition to radiation damping. For situations where αL is sufficiently large, leaky wave amplitudes may be estimated from a single ray contribution instead of including a superposition of many contributions leading to resonances.[9,11] When that is not the case, the results of Secs. IV and V may be easily extended to give the desired superposition. Anisotropy in the leaky wave propagation parameters (neglected in this analysis) may need to be included for description of helical wave resonances of some elastic targets.[35]

ACKNOWLEDGMENTS

APPENDIX A: LEAKY WAVE RESIDUES FOR NONDISSIPATIVE SURFACES

For the cases where the plane surface is elastic (and thus nondissipative), $R(\zeta)$ is unimodular for real horizontal wave numbers ζ larger than a critical value ζ_S, the shear critical value in the case of an elastic half space. For a vacuum-backed truly elastic plate, R is unimodular for all real ζ. It is convenient to use the following factorization[25,29] that satisfies the symmetry requirement $R(-\zeta) = R(\zeta)$ where ζ may be complex:

$$R(\zeta) = R_0(\zeta) \left(\frac{\zeta^2 - k_o^2}{\zeta^2 - k_p^2} \right) \tag{A1}$$

$$= R_0(\zeta) \left(1 + \frac{k_p^2 - k_o^2}{\zeta^2 - k_p^2} \right), \tag{A2}$$

where $R_0(\zeta) = R_0(-\zeta)$ is regular near the poles of R at $\zeta = \pm k_p$ and $k_p = k_l + i\alpha$. For real ζ giving $|R| = 1$, R_0 is unimodular provided the location of the complex zeros at $\pm k_o$ are such that $k_o = k_p^* = k_l - i\alpha$. That is the situation for the elastic half space and the vacuum-backed plate. It follows from $k_o = k_p^*$ that without any approximation:

$$R_0(k_l) = -R(k_l) \left(\frac{2k_l + i\alpha}{2k_l - i\alpha} \right) = -R(k_l) e^{i2 \tan^{-1}(\alpha/2k_l)}, \tag{A3}$$

$$\text{Res}(R, k_p) = 2i\alpha(k_l/k_p) R_0(k_p). \tag{A4}$$

For weak fluid loading, k_p is close to k_l giving

$$\text{Res}(R, k_p) \approx 2i\alpha(k_l/k_p) R_0(k_l) = 2i\alpha(k_l/k_p) e^{i\varphi_{bl}}, \tag{A5}$$

where $\varphi_{bl} = \arg(R_0(k_l))$ for the lth leaky pole under consideration. For example, leaky Rayleigh waves on (i) stainless-steel–water and (ii) aluminum–water interfaces have α/k_l of (i) 0.011 and (ii) 0.029. To test the approximation in (A5), the following quantities were numerically computed: $f_1 \equiv \text{Res}(R, k_p)/2i\alpha$, $f_2 \equiv (k_l/k_p) R_0(k_l)$, and φ_{bl}. The results are: (i) $|f_1| = 1.003$, $|f_2| = 0.999$, $\arg(f_1) = 0.10$, $\arg(f_2) = 0.10$, $\varphi_{bl} = 0.10$; (ii) $|f_1| = 1.026$, $|f_2| = 0.999$, $\arg(f_1) = 0.29$, $\arg(f_2) = 0.30$; $\varphi_{bl} = 0.30$. These results support the use of (A5) when $\alpha/k_l \ll 1$. From (A3), a phase factor $\exp[i2 \tan^{-1}(\alpha/2k_l) - i \tan^{-1}(\alpha/k_l)] \approx \exp[i(\alpha/k_l)^3/4]$ which is usually negligible may be factored from (A5). It follows that $\arg(f_1) \approx \arg(-R(k_l))$. Unless ka is very large, pole locations may need to be found from Watson methodology[9,16] or as discussed in Ref. 27. Pole locations for leaky Rayleigh waves on curved solids appear to converge at large ka to the corresponding flat surface results as evidenced by numerical results in Ref. 9 and of other authors.[39]

For leaky Lamb waves the phase of $-R(k_l)$, and thus also φ_{bl} for the lth pole from (A3), can be significantly affected by the poles for the other Lamb waves.[14] For the s_0 wave at sufficiently low frequencies, it can be useful to approximate R_0 by the reflection coefficient of a vacuum-backed infinitesimally thin plate of thickness h having the same mass per area $h\rho_E$ as the plate or shell of interest. This approximation gives[18]

$$\varphi_{bl} = -2 \tan^{-1}[\rho/(kh\rho_E \cos \theta_l)]. \tag{A6}$$

The value of $-R(k_l)$ when $l = a_0$ is offset by the s_0 pole (which has a smaller value of k_l) so that $|\varphi_{bl}|$ is much smaller than in (A6). This has been demonstrated by plotting $\arg(R(k \sin \theta))$ as a function of θ and is consistent with direct numerical evaluation of the phase of the far-field coupling coefficient G_l (see Appendix B) from the Watson transformation for a thick spherical shell.[14] Pole evolution for shells is discussed, e.g., in Refs. 10–13.

For Lamb waves on a plate with symmetric fluid loading, $R(\zeta)$ is no longer unimodular for real ζ so that the forms of (A2)–(A4) are altered. It is possible for the k_o to lie on the real axis such that numerical evaluation[27] gives $k_o \approx k_l$. With this approximation, (A1) gives

$$\text{Res}(R, k_p) \approx i\alpha R_0(k_p), \tag{A7}$$

which has approximately half the magnitude of (A4) when other factors are not significantly altered. This follows from

other approximations and from symmetry.[18] For both symmetric and unsymmetric fluid loading the numerator in the second term of (A2) is regular and nonvanishing near $\zeta = k_p$ leading to the form used in Eq. (1).

APPENDIX B: UNCOATED CYLINDER VIA EQ. (9) AND FAR-FIELD COUPLING COEFFICIENT

For the uncoated cylinder $z_E \to -\infty$ and the y integration in (16) is evaluated using (9). It is convenient to introduce the relative angle $\bar\theta = (x'/a) - \theta_l$; (16) gives

$$p_l \approx i \; \mathrm{Res}(R, k_p) a p_B e^{i\varphi_{BS}} I, \tag{B1a}$$

$$I = \int_{\bar\theta_L}^{\bar\theta_U} \exp(i\varphi) \exp(a\alpha\bar\theta) d\bar\theta, \tag{B1b}$$

$$\varphi = ka \sin\theta_l \sin\bar\theta + ka \cos\theta_l (1 - \cos\bar\theta) - k_l a \bar\theta$$

$$\approx (ka\bar\theta^2/2)\cos\theta_l, \tag{B1c}$$

where $\varphi_{BS} = (k_l + i\alpha)a(\theta_S - \theta_l)$ and $\bar\theta_{U,L} = \theta_{U,L} - \theta_l$ with $\theta_{U,L} = \pm \pi/2$. These are analogous to Eqs. (M11)–(M14) where they are discussed in more detail. The following evaluation of I uses the quadratic approximation of φ noted in (17c). The result is[36]

$$I = \Delta\theta_f \exp(i\pi/4)\exp(ib^2)(F_L - F_U), \tag{B2}$$

where $F_{U,L} = F(\xi_{U,L} - ib)$ denote Fresnel integrals of complex argument (24d), $\Delta\theta_f$ and b are given by (17), and $\xi_{U,L} = \pi^{1/2}(\theta_{U,L} - \theta_l)/\Delta\theta_f$. Approximations for F with large $|\xi|$ given in (25) are such that in the high-frequency limit of small $\Delta\theta_f$, $F_L \to 1$ and $F_U \to 0$ and $I \approx \Delta\theta_f \exp(i\pi/4)\exp(ib^2)$ as previously discussed.[18] Application of the approximate residue (A5) yields a prefactor in (B1a) like that of Eq. (M11), where previously $R_0(k_l)$ was denoted by the phase factor $\exp(i\varphi_{bl})$ and (k_l/k_p) was omitted.

For the purpose of illustrating the small phase errors introduced, it is helpful to review the consequences of geometric propagation of the radiation by the leaky wave in the vicinity of surface point S. Flux conservation within a ray tube diverging from a virtual line source behind S (shown at F in Fig. 6 for backscattering) yields the following far-field amplitude at a distance $r_C \gg ka^2$ from the center of the cylinder[18]

$$p_l(s, r_C) \approx -(G/2) p_C (a/2r_C)^{1/2} e^{i\eta} e^{-\beta(\theta_S - \theta_l)} e^{ikr_C}, \tag{B3}$$

where $\beta = \alpha a$ and the phase conventions are as discussed in M Sec. II. Then Eq. (B2) yields the following approximation for the coupling coefficient:

$$G_l = 4(\pi a/k)^{1/2} e^{-i\pi/4} \; \mathrm{Res}(R, k_p)(F_L - F_U) e^{ib^2}, \tag{B4a}$$

$$\approx [8\pi\beta e^{i\pi/4}/(\pi ka)^{1/2}]\{R_0(k_l)[1 + i(\alpha/k_l)]^{-1}$$

$$\times (F_L - F_U)\} e^{ib^2}, \tag{B4b}$$

where (B4b) uses (A5). The phase factor $\exp(ib^2)$ may be attributed to a shift of the apparent ray path through B to a path through $\theta > \theta_l$ because of the damping (M, Sec. III). It is usually negligible for most leaky wave parameters at high

frequencies where $(F_L - F_U) \approx 1$. It is plausible that the factor $(F_L - F_U)$ in (B2) and (B4) exaggerates the effects of the shadow boundaries at L and U in Fig. 6 except when ka is very large. Neglecting an $O(\alpha/ka)^2$ magnitude correction and taking $|R_0(k_l)| = 1$ shows the resulting $|G_l|$ reduces to the approximation from Refs. 3, 9, 11, and 15 and from the ''pure ray theory'' in Ref. 7, all derived using different formulations than Eq. (2) used here. Numerical calculations of $\mathrm{Res}(R, k_p)$ summarized in Appendix A for leaky (flat surface) Rayleigh waves indicate the factor $R(k_l)[1 + i(\alpha/k_l)]^{-1}$ is close to unity with an approximately $O(\alpha/k_l)$ phase. That result indicates the high-frequency behavior for solid cylinders. For empty shells the phase behavior is more complicated and depends on the type of leaky Lamb wave under consideration as summarized in Appendix A. The phase $\varphi_{bl} = \arg R_0(k_l)$ corresponds to the high-frequency phase correction given in Eq. (21b) of Ref. 15 (derived by the method of Ref. 9). For the appropriate case (see Appendix A) φ_{bl} becomes the ''pure ray'' result of Eq. (4) of Ref. 7. The aforementioned limiting approximation for G also gives agreement between (B3) and the broadside limit of Eq. (31) which has $\psi = \pi/2$.

There can be additional $O(1/ka)$ phase corrections that do not appear to be as easily included in the present formulation. [These correspond to the equivalent small phase corrections given by Eq. (11a) of Ref. 4 and (22b) of Ref. 15. They are also evident in the difference between the curves for the numerically evaluated residues from thin shell theory and the ''pure ray'' results given in Fig. 1(b) of Ref. 7.] For reasons explained in the Introduction, these $O(1/ka)$ phase corrections may be unimportant for many applications of the three-dimensional formulation given by Eq. (2). These phase corrections appear to be wave kinematic in nature, as evidenced by the difference in phase shifts calculated for diverging or converging waves in comparison to plane-wave phase shifts. Such phase modifications were not included in the far-field propagation of (B1) by ray theory leading to Eqs. (B3) and (B4).

The case of a θ-coated cylinder (Fig. M2) is recovered by replacing θ_L by the coating angle θ_D in (B2). For a cylindrical surface of variable curvature at normal incidence (Fig. M4), Eq. (16) with $y_E \to -\infty$ reduces to a one-dimensional integral by using (9). The result reduces to Eq. (M8) when the aforementioned approximation to (A5) is used and limiting values for x_U and x_L are inserted.

[1] G. Kaduchak, C. M. Wassmuth, and C. M. Loeffler, ''Elastic wave contributions in high resolution acoustic images of fluid-filled finite cylindrical shells in water,'' J. Acoust. Soc. Am. **100**, 64–71 (1996); S. S. Dodd, ''Sonar imaging of elastic fluid-filled cylindrical shells,'' Ph.D. thesis, University of Texas at Austin, 1995.

[2] T. L. Henderson, ''Dynamic stereo imaging of hard-skinned sonar targets,'' J. Acoust. Soc. Am. **96**, 3311 (1994); B. L. Douglas, J. M. Silkaitis, and H. Lee, ''Single-reference calibration technique for multiple-element array sonar imaging systems,'' *ibid.* **100**, 400–407 (1996); J. L. Sutton, ''Underwater acoustic imaging,'' Proc. IEEE **67**, 554–566 (1979). For imaging of relatively low ka processes see, e.g., E. G. Williams, ''Imaging the sources on a cylindrical shell from far-field pressure measured on a semicircle,'' J. Acoust. Soc. Am. **99**, 2022–2032 (1996).

[3] L. B. Felsen, J. M. Ho, and I. T. Lu, ''Three-dimensional Green's function for fluid-loaded thin elastic cylindrical shell: Alternative representations

and ray acoustic forms,'' J. Acoust. Soc. Am. **87**, 554–569 (1990); Erratum **89**, 1463–1464 (1991).

[4] J. M. Ho, ''Acoustic scattering by submerged elastic cylindrical shells: Uniform ray asymptotics,'' J. Acoust. Soc. Am. **94**, 2936–2946 (1993).

[5] M. L. Rumerman, ''Application of the Sommerfeld–Watson transformation to scattering of acoustic waves obliquely incident upon cylindrical shells,'' J. Acoust. Soc. Am. **91**, 2503–2509 (1992).

[6] A. N. Norris and D. A. Rebinsky, ''Acoustic coupling to membrane waves on elastic shells,'' J. Acoust. Soc. Am. **95**, 1809–1829 (1994).

[7] D. A. Rebinsky and A. N. Norris, ''Benchmarking an acoustic coupling theory for elastic shells of arbitrary shape,'' J. Acoust. Soc. Am. **98**, 2368–2371 (1995).

[8] Y. Yang, A. N. Norris, and L. S. Couchman, ''Acoustic scattering from fluid-loaded elastic shells: A Gaussian beam approach,'' J. Acoust. Soc. Am. **98**, 611–622 (1995).

[9] 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).

[10] S. G. Kargl and P. L. Marston, ''Observations and modeling of the backscattering of short tone bursts from a spherical shell: Lamb wave echoes, glory, and axial reverberations,'' J. Acoust. Soc. Am. **85**, 1014–1028 (1989); Erratum **89**, 2462 (1991).

[11] N. H. Sun and P. L. Marston, ''Ray synthesis of leaky Lamb wave contributions to backscattering from thick cylindrical shells,'' J. Acoust. Soc. Am. **91**, 1398–1402 (1992).

[12] L. G. Zhang, N. H. Sun, and P. L. Marston, ''Midfrequency enhancement of the backscattering of tone bursts by thin spherical shells,'' J. Acoust. Soc. Am. **91**, 1862–1874 (1992); G. Kaduchak and P. L. Marston, ''Observation of the midfrequency enhancement of tone bursts backscattered by a thin spherical shell in water near the coincidence frequency,'' *ibid.* **93**, 224–230 (1993).

[13] G. Kaduchak, D. H. Hughes, and P. L. Marston, ''Enhancement of the backscattering of high-frequency tone bursts by thin spherical shells associated with a backwards wave: Observations and ray approximation,'' J. Acoust. Soc. Am. **96**, 3704–3714 (1994).

[14] S. G. Kargl and P. L. Marston, ''Background contributions and coupling coefficients for backscattering by thick shells,'' J. Acoust. Soc. Am. **101**, 3792–3797 (1997).

[15] P. L. Marston, ''Variable phase coupling coefficient for leaky waves on spheres and cylinders from resonance scattering theory,'' Wave Motion **22**, 65–74 (1995).

[16] K. L. Williams and P. L. Marston, ''Backscattering from an elastic sphere: Sommerfeld–Watson transformation and experimental confirmation,'' J. Acoust. Soc. Am. **78**, 1093–1102 (1985); Erratum **79**, 2091 (1986).

[17] P. L. Marston, ''Leaky waves on curved scatterers. I. Fresnel width of coupling regions and elliptical Fresnel patches,'' J. Acoust. Soc. Am. **96**, 1893–1898 (1994). See also applications of Fermat's principle in: J. B. Keller and F. C. Karal, ''Surface wave excitation and propagation,'' J. Appl. Phys. **31**, 1039–1046 (1960).

[18] P. L. Marston, ''Leaky waves on weakly curved scatterers. II. Convolution formulation for two-dimensional high-frequency scattering,'' J. Acoust. Soc. Am. **97**, 34–41 (1995). A minus sign was omitted from an expression for φ_b given below (A2).

[19] P. L. Marston, ''Approximate meridional leaky ray amplitudes for tilted cylinders: end-backscattering enhancements and comparisons with exact theory for infinite solid cylinders,'' J. Acoust. Soc. Am. **101**, 358–369 (1997); ''Approximation for leaky wave amplitudes in acoustic imaging: applications to high frequency sonar,'' Proceedings of the 23rd International Symposium on Acoustical Imaging, edited by S. Lees (submitted). In the first reference an overall factor of 2 was omitted from the right side of Eq. (12).

[20] N. G. Van Kampen, ''The method of stationary phase and the method of Fresnel zones,'' Physica (Amsterdam) **24**, 437–444 (1958).

[21] Y. A. Kravtsov and Y. I. Orlov, ''Limits of applicability of the method of geometric optics and related problems,'' Sov. Phys. Usp. **23**, 750–762 (1980).

[22] P. L. Marston, ''Geometrical and catastrophe optics methods in scattering,'' in *Physical Acoustics*, edited by R. N. Thurston and A. D. Pierce (Academic, Boston, 1992), Vol. 21, pp. 1–234, Secs. 2.5 and 2.14.

[23] M. Conti, ''Mid-frequency acoustic scattering from finite internally-loaded cylindrical shells near axial incidence,'' Ph.D. thesis, Massachusetts Institute of Technology, 1995.

[24] D. T. DiPerna and T. K. Stanton, ''Fresnel zone effects in the scattering of sound by cylinders of various lengths,'' J. Acoust. Soc. Am. **90**, 3348–3355 (1991).

[25] H. L. Bertoni and T. Tamir, ''Unified theory of Rayleigh-Angle phenomena for acoustic beams at liquid-solid interfaces,'' Appl. Phys. **2**, 157–172 (1973).

[26] D. E. Chimenti, J. G. Zhang, S. Zeroug, and L. B. Felsen, ''Interaction of acoustic beams with fluid-loaded elastic structures,'' J. Acoust. Soc. Am. **95**, 45–59 (1994).

[27] S. Zeroug and L. B. Felsen, ''Nonspecular reflection of two- and three-dimensional acoustic beams from fluid-immersed plane-layered elastic structures,'' J. Acoust. Soc. Am. **95**, 3075–3089 (1994); ''Nonspecular reflection of two- and there-dimensional acoustic beams from fluid-immersed cylindrically layered elastic structures,'' J. Acoust. Soc. Am. **98**, 584–598 (1995).

[28] R. H. Tew, ''Non-specular reflection phenomena at a fluid-solid boundary,'' Proc. R. Soc. London, Ser. A **437**, 433–449 (1992); ''The excitation and propagation of a two-dimensional ''leaky '' Rayleigh wave on an elastic solid under compressible fluid loading,'' Wave Motion **22**, 17–29 (1995).

[29] H. L. Bertoni, ''Ray-optical evaluation of $V(z)$ in the reflection acoustic microscope,'' IEEE Trans. Sonics Ultrason. **SU-31**, 105–116 (1984).

[30] P. Zinin, W. Weise, O. Lobkis, O. Kolosov, and S. Boseck, ''Fourier optics analysis of spherical particles image formation in reflection acoustic microscopy,'' Optik (Stuttgart) **98**, 45–59 (1994).

[31] G. V. Frisk, *Ocean and Seabed Acoustics: A Theory of Wave Propagation* (Prentice-Hall Englewood, Cliffs, NJ, 1994), Sec. 4.8.

[32] B. Noble, *Methods Based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations* (Chelsea, New York, 1988), Sec. 2.5.

[33] J. J. Stamnes, ''Role of backward waves in backscattering of ultrasonic beams at a liquid–solid interface,'' J. Acoust. Soc. Am. **76**, 627–631 (1984).

[34] M. Rousseau and Ph. Gatignol, ''Short wave analysis for the reflection of bounded acoustic beams onto liquid–solid interfaces at the Rayleigh incidences,'' J. Acoust. Soc. Am. **78**, 1859–1867 (1985).

[35] A. D. Pierce and H.-G. Kil, ''Elastic wave propagation from point excitations on thin-walled cylindrical shells,'' J. Vib. Acoust. (Trans. ASME) **112**, 399–406 (1990).

[36] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965), Eq. (7.4.32).

[37] G. A. Deschamps, S. W. Lee, E. Gowan III, and T. Fontana, ''Diffraction of an evanescent plane wave by a half plane,'' Wave Motion **1**, 25–35 (1979).

[38] X. L. Bao, ''Echoes and helical surface waves on a finite elastic cylinder excited by sound pulses in water,'' J. Acoust. Soc. Am. **94**, 1461–1466 (1993).

[39] G. V. Frisk, J. W. Dickey, and H. Überall, ''Surface wave modes on elastic cylinders,'' J. Acoust. Soc. Am. **58**, 996–1008 (1975).

Aeroelastic structural acoustic coupling: Implications on the control of turbulent boundary-layer noise transmission

Robert L. Clark and Kenneth D. Frampton

Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708-0300

A method of formulating a model to evaluate the aeroelastic structural acoustic response of a panel subjected to turbulent boundary layer (TBL) noise sources and coupled with full potential flow aerodynamics is presented. Reduced-order models of both the aerodynamics and the structural acoustic coupling are presented such that a state-variable realization of the entire system dynamics can be developed for future active control system design and synthesis with modern and robust control theory. Results from this study demonstrate the importance of including aeroelastic coupling in modeling the structural acoustic response of panels for interior noise control on modern aircraft. At subsonic flow conditions, the aeroelastic coupling serves to increase the transmission loss across the panel with increasing Mach number; however, the power spectrum of the TBL noise source increases with increasing Mach number as well and thus offsets this benefit to some degree. Results from this study also serve to demonstrate that for future analysis of robust stability and performance, variations in the plant dynamics due to variations in flow conditions must be considered in the design of broadband, feedback, active structural acoustic control systems. © *1997 Acoustical Society of America.* [S0001-4966(97)00409-8]

PACS numbers: 43.40.Vn [PJR]

LIST OF SYMBOLS

A,B,C,D	state space matrices
a	panel length (chord)
a_a	external speed of sound
b	panel width (span)
D	$EI/(1-\nu^2)$, stiffness
D_{mn}	aerodynamic influence coefficient
$H_{mn}(t)$	aerodynamic influence function
h_s	plate thickness
$I_{mn}(t)$	aerodynamic influence function
M	Mach number
M_n	mass matrix
N	number of plate expansion modes
$p^a(x,y,z,t)$	aerodynamic pressure
$p^{tbl}(x,y,z,t)$	TBL pressure
$q_n(t)$	plate generalized coordinate
$Q_n^a(t)$	aerodynamic generalized force
$Q_n^c(t)$	cavity generalized force
$Q_n^{tbl}(t)$	TBL generalized force
s	tU_a/a
S	plate surface area
S_{mn}	aerodynamic influence coefficient
t	time
U_a	free stream velocity
$w(x,y,t)$	panel displacement
x,y,z	Cartesian coordinates
Λ_m	cavity modal volume
$\Psi_n(x,y)$	plate modal expansion functions
$\phi^a(x,y,z,t)$	aerodynamic velocity potential
ρ_a	external fluid density
ρ_s	plate density

INTRODUCTION

Reduction in the transmission of turbulent boundary layer (TBL) noise to the interior fuselage of modern business and commercial jet aircraft operating under cruise conditions is gaining increasing concern in an ever competitive market to reduce overall interior noise levels. Unlike the periodic noise sources typical of propeller driven aircraft, TBL noise is of stochastic nature and thus adaptive feedforward control strategies are not practical in the absence of an uncontrollable reference signal. To this end, Thomas and Nelson (1995) investigated linear quadratic regulator (LQR) theory and linear quadratic Gaussian (LQG) methods to design a feedback controller for a panel subjected to TBL noise on one surface and coupled to an acoustic half-space on the

opposite surface. The controller was designed based upon the augmented system dynamics as outlined by Baumann *et al.* (1991), which ensures that the structural acoustic response (as opposed to the structural response) is minimized in the controller optimization. The critical element in the realization of such control systems is the design and implementation of radiation filters which capture the structural acoustic coupling.

Thomas and Nelson (1995) extended the work by including a simplified model of the TBL noise as outlined by Corcos (1967). The frequency-dependent characteristics of the TBL noise was captured in the realization of filters through spectral factorization, similar to that performed to obtain the radiation filters. These TBL noise filters were then placed in series with the panel (no feedback in the dynamic

a) Schematic of Panel in Baffle

b) Profile of Panel in Baffle
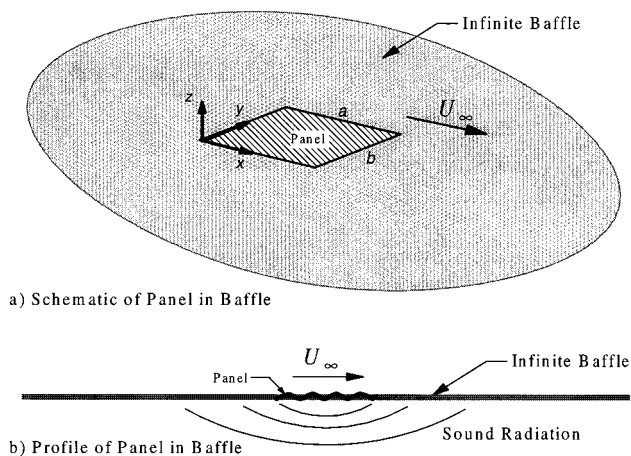
FIG. 1. Schematic diagram of the system under consideration.

FIG. 2. Block diagram of coupled aeroelastic-structural acoustic system.

model) and served to provide a more realistic measure of the bandwidth of the disturbance source. Hence, the TBL noise filters ensure that the control effort is optimized for the bandwidth of the disturbance source (shaped by the TBL noise filters). Preliminary results from this study showed that feedback control demonstrates potential in the control of sound transmission across exterior panels of the fuselage. However, within this study, the effects of the mean-flow on the panel dynamics were neglected.

The objective of the work outlined herein is to demonstrate the importance of considering aeroelasticity in the structural acoustic model for active control applications. Aeroelasticity is analogous to feedback control in the sense that the structural modes are coupled through the dynamics of the mean flow. As will be demonstrated, the panel dynamics vary as a function of flight conditions, and thus any feedback control system designed for increasing the transmission loss across a panel subjected to mean flow must include the dynamics associated with the mean flow. In recent research, Romanowski and Dowell (1995), as well as Hall *et al.* (1995), have shown that it is possible to develop reduced order models of fluid systems and cast those models in terms of an eigenvalue problem. Frampton *et al.* (1996a) extended this work to include the reduced order modeling of panels subjected to mean flow and cast the models in state-variable form for the purpose of control system design and synthesis. More recent work by Frampton and Clark (1996) demonstrates that singular value decomposition (SVD) can be used to further reduce the order of a full potential flow, aeroelastic model, rendering a system amenable to analysis and design with typical modern and robust control strategies. The development of such models will be reviewed in this work along with results demonstrating the effects of aeroelastic coupling on the structural acoustic response of a panel radiating into a half-space.

I. MODEL DEVELOPMENT

A schematic diagram of the system is illustrated in Fig. 1. As depicted in Fig. 1(a), the panel is positioned within an infinite baffle, and air with mean velocity U_∞ flows over the
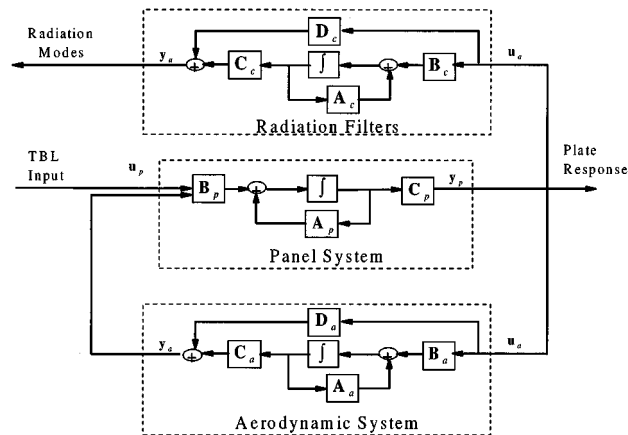
surface of the panel in the x direction. The panel has dimensions a and b in the x and y directions, respectively, with out-of-plane displacement occurring in the z direction. As illustrated in Fig. 1(b), the opposite side of the panel is subject to a zero mean velocity (no flow) condition and the sound power radiated to this half-space is computed through the implementation of radiation filters derived from the Rayleigh integral.

A block diagram of the coupled system is illustrated in Fig. 2. Central to the block diagram are the panel dynamics, which in this study were obtained from an assumed modes approach (Meirovitch, 1967). As illustrated, the output of the panel dynamics are coupled through the aerodynamic system which constitutes the aeroelastic response of the panel. This feedback through the fluid couples the panel modes and can potentially lead to divergence at Mach numbers less than one or flutter at Mach numbers greater than one in extreme cases. Of course divergence or flutter is of no concern for panels designed for operation on modern aircraft. However, the dynamic coupling of the panel to the fluid influences the sound power radiated to the interior as will be shown and thus affects feedback control system design.

A TBL noise source also acts on the panel and serves as the disturbance. Consistent with the assumptions made by Thomas and Nelson (1995), the cross-coupling is ignored in this simple model. However, one should recognize that cross-coupling of modes through the fluid dynamics is captured through the aeroelastic response of the system. The output of the panel is then passed to a state-variable model of the radiation filters such that the rms (root-mean-square) sound power can be computed. Spectral factorization is used to develop state-variable models of the radiation filters and the TBL noise filters.

Thus, the complete system can be divided into four sections: (1) panel model, (2) TBL noise model, (3) aerodynamic model, and (4) radiation filter model. Each of these models will be discussed in the following subsections.

A. Panel model

The partial differential equation of motion describing a thin, uniform panel including external fluid loading and TBL loading can be expressed as follows (Meirovitch, 1967):

$$\rho_s h_s \frac{\partial^2 w(x,y,t)}{\partial t^2} + D\left[\frac{\partial^4 w(x,y,t)}{\partial x^4} + \frac{\partial^4 w(x,y,t)}{\partial x^2 \, \partial y^2}\right.$$

$$\left. + \frac{\partial^4 w(x,y,t)}{\partial y^4} + p^a(x,y,z=0,t)\right.$$

$$\left. + p^{\text{tbl}}(x,y,z=0,t) = 0. \right. \tag{1}$$

The coordinate system for the panel and associated flow was illustrated in Fig. 1.

A separable expansion solution is assumed using the *in vacuo* orthogonal panel eigenfunctions and generalized coordinates of the form:

$$w(x,y,t) = \sum_{n=1}^{N} \Psi_n(x,y) q_n(t). \tag{2}$$

Substituting Eq. (2) into Eq. (1), multiplying by an arbitrary expansion function, $\Psi_m(x,y)$, and integrating over the domain yields a set of ordinary differential equations of the form (Meirovitch, 1967)

$$M_n[\ddot{q}_n(t) + \omega_n^2 q_n(t)] + \rho_a U_a^2 Q_n^a(t) + Q_n^{\text{tbl}}(t) = 0, \tag{3}$$

where the aerodynamic generalized forces are

$$Q_n^a(t) = \int_0^{b'} \int_0^a \frac{p^a(x,y,z=0,t)}{\rho_a U_a^2} \Psi_n(x,y) dx \, dy, \tag{4}$$

and the turbulent boundary layer generalized forces are

$$Q_n^{\text{tbl}} = \int_0^b \int_0^a p^{\text{tbl}}(x,y,z=0,t) \Psi_n(x,y) dx \, dy. \tag{5}$$

A structural proportional damping term, $2\eta_s \omega_n \dot{q}_n(t)$, was also introduced into Eq. (3) where η_s is the damping ratio. An additional term can appear in Eq. (3) as a result of piezoelectric actuation. Details concerning piezoelectric transduction can be found in the work by Frampton *et al.* (1996b).

B. TBL noise model

The disturbance of interest in this system is a result of turbulent boundary layer (TBL) pressures on the surface of the plate. Beginning with the the fundamental equation for the TBL generalized forces on the plate stated in Eq. (5) and employing Fourier transform and power spectrum relations, the generalized TBL force power spectrum can be expressed as (Corcos, 1967; Bull, 1996)

$$\mathscr{P}_{Q_m Q_n}(\omega) = \int_S \int_{S^*} \mathscr{P}_{pp}(\omega;x,y,x^*,y^*) \Psi_m(x,y)$$

$$\times \Psi_n(x^*,y^*) dS \, dS^*. \tag{6}$$

The challenge presented with this formulation is to obtain an accurate representation of the TBL spatial/temporal correlation function $\mathscr{P}_{pp}(\omega;x,y,x^*,y^*)$. The form used here is consistent with that presented by Corcos (1967):

$$\mathscr{P}_{pp}(\omega;x,y,x^*,y^*) = \mathscr{S}_{pp}(\omega)\exp\left[-\left|\frac{\omega(x-x^*)}{\alpha_1 U_c}\right|\right]$$

$$\times \exp\left[-\left|\frac{\omega(y-y^*)}{\alpha_2 U_c}\right|\right]$$

$$\times \exp\left[\frac{-i\omega(x-x^*)}{U_c}\right]. \tag{7}$$

For the frequency range of interest ($f < 1000$ Hz), the frequency-dependent function, $\mathscr{S}_{pp}(\omega)$, is approximately constant in frequency and scales like $(\rho_a U_a^2)^2$ in amplitude (Blake, 1986). The constants α_1 and α_2 control the spatial correlation of the TBL pressures and have been empirically determined to be approximately 10 and 2, respectively (Thomas and Nelson, 1995). A boundary layer convection velocity of $U_c = 0.6 \times U_\infty$ was used.

As discussed earlier in this section, a further approximation used in this investigation was to assume that the cross-modal power spectra, i.e., $\mathscr{P}_{Q_m Q_n}$ for $m \neq n$, is zero. This is equivalent to neglecting the plate modal coupling which results from the TBL model. This assumption is justified since the cross-coupling terms ($m \neq n$) are much less than the mutual terms ($m = n$) due to the relatively large separation between the plates resonant peaks.

It is well known that the Corcos TBL model is only accurate in the region of the convective ridge and is inaccurate for low wave number plate motion (Bull, 1996). For this system the plate motion is proximal to the TBL convective ridge in the low Mach number region. The advantage of the Corcos TBL model is that there is a closed form solution for the generalized force power spectra which results in a large computational savings. A further assumption made is that the TBL pressures are independent of the plate motion. The emphasis of this work is not placed upon advancing TBL noise models, but rather in formulating a system model which provides all the necessary physics to begin preliminary evaluation of active structural acoustic control in the presence of aeroelastic coupling.

The TBL power spectra, $\mathscr{P}_{Q_m Q_n}(\omega)$, for each of the first four structural modes used in this model are shown in Fig. 3(a) and (b) and were found as outlined by Thomas and Nelson (1995). For the purpose of comparison, the TBL power spectra for both subsonic ($M = 0.5$) and supersonic ($M = 1.5$) flow are illustrated. As previously noted, the TBL pressures scale with the square of the flow velocity, and the shape of the spectrum effectively ''stretches'' with increasing flow speed. In addition, a filter obtained through spectral factorization can be used to represent this power spectrum (Thomas and Nelson, 1995). A filter was designed in this study and placed in series with the dynamic model of the system to ''shape'' the spectral content of the input disturbance. The shape of this filter serves an important role in defining the bandwidth where the most control energy is required.

C. Aerodynamic model

Dugundji (1966) published a review of early investigations into linear panel flutter and more recently Gray and

a) TBL power spectrum for M = 0.5

b) TBL power spectrum for M = 1.5
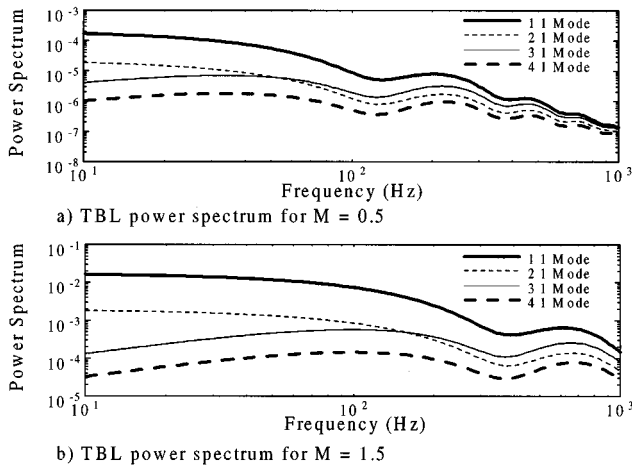
FIG. 3. Representative TBL power spectra.

Mei (1992) reviewed recent efforts including nonlinear effects. However, the majority of the work to date has been performed using ''piston theory'' as the aerodynamic model (Hajela, 1991; Scott, 1994). The use of piston theory is often practical since at higher Mach numbers the method proves reasonably accurate, and modeling the effects of potential flow aerodynamics at lower Mach numbers ($M < 1.5$) is challenging. Most of the references which include linearized potential flow aerodynamics focus on the determination of flutter boundaries and subsequently the limit cycle oscillations in the postflutter (nonlinear) regime (Dowell, 1967, 1975). These models are not conducive to the application of modern or robust control theory since a state-variable realization of the dynamics is most convenient. More recent efforts by Frampton *et al.* (1996a) and Frampton and Clark (1996) have been devoted to the development of reduced-order models of aerodynamics associated with full potential flow theory cast in state-variable form. A brief overview of this method is outlined herein to convey the essential elements required in this study.

1. Transform solution to the linearized potential flow generalized forces

Consider the case of a simply supported panel with the following eigenfunctions:

$$\Psi_n(x,y) = \sin\left(\frac{n\pi}{a} x\right) \sin\left(\frac{\pi}{b} y\right).$$ (8)

Note that only one mode in the y direction is included such that $n = 1,2,...,N$. This assumption is made in order to simplify the model and because the chordwise modes are most significantly influenced by the external flow. In future work with piezoelectric actuation more lateral modes will be included in the assumed solution.

The aerodynamic generalized forces, $Q_n^a(t)$, are found by solving the partial differential equation describing the velocity potential in an inviscid, irrotational fluid moving parallel to the x axis

$$\nabla^2 \phi^a - \frac{1}{a_\infty^2}\left[\frac{\partial^2 \phi^a}{\partial t^2} + 2U_a \frac{\partial^2 \phi^a}{\partial x \, \partial t} + U_a^2 \frac{\partial^2 \phi^a}{\partial x^2}\right] = 0$$ (9)

subject to the boundary conditions for a panel embedded in an infinite baffle

$$\left.\frac{\partial \phi^a}{\partial z}\right|_{z=0} = \begin{cases} \dfrac{\partial w}{\partial t} + U_a \dfrac{\partial w}{\partial x} & \text{on the plate,} \\ 0 & \text{off the plate,} \end{cases}$$ (10)

and a finiteness or radiation condition as z approaches infinity.

The solution can be obtained by taking a Laplace transform with respect to time and a double Fourier transform with respect to the x and y spatial coordinates. The transformation is applied to Eqs. (9) and (10) and Bernoulli's equation

$$p^a = -\rho_a\left[\frac{\partial \phi^a}{\partial t} + U_a \frac{\partial \phi^a}{\partial x}\right],$$ (11)

while incorporating Eq. (2). Details of the solution are discussed by Dowell (1967, 1975).

The solution yields the generalized aerodynamic forces on the panel which are given here as

$$Q_n^a(t) = \sum_{m=1}^{N} Q_{mn}^a(t),$$ (12)

where $Q_{mn}^a(t)$ is the force on the nth panel mode due to motion of the mth panel mode such that

$$Q_{mn}^a(t) = q_m(t)S_{mn} + \dot{q}_m(t)D_{mn}$$ (13)

$$+ \int_0^t q_m(\tau)H_{mn}(t-\tau)d\tau$$

$$+ \int_0^t \dot{q}_m(\tau)I_{mn}(t-\tau)d\tau.$$ (14)

where

$$S_{mn} = \frac{1}{M} \int_0^a \int_0^b \frac{\partial \Psi_m}{\partial x} \Psi_n \, dy \, dx,$$ (15)

$$D_{mn} = \frac{1}{MU_a} \int_0^a \int_0^b \Psi_m \Psi_n \, dy \, dx,$$ (16)

$$H_{mn}(t) = -\frac{U_a}{4\pi^2 M^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{mn} i\alpha\sqrt{\alpha^2+\gamma^2} e^{-i\alpha U_a t}$$

$$\times J_1[a_\infty t \sqrt{\alpha^2+\gamma^2}]d\alpha \, d\gamma,$$ (17)

$$I_{mn}(t) = -\frac{1}{4\pi^2 M^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_{mn} \sqrt{\alpha^2+\gamma^2} e^{-i\alpha U_a t}$$

$$\times J_1[a_\infty t \sqrt{\alpha^2+\gamma^2}]d\alpha \, d\gamma,$$ (18)

and

$$G_{mn} = \int_0^a \int_0^b \Psi_m(x,y)e^{-i(\alpha x+\gamma y)}dy \, dx$$

$$\times \int_0^a \int_0^b \Psi_n(x,y)e^{i(\alpha x+\gamma y)}dy \, dx.$$ (19)

It is important to note that this solution of the linearized potential flow equation is inaccurate in the transonic flow

region where nonlinear effects become important. More information on this modeling technique may be found in the references by Frampton *et al.* (1996a) and Dowell (1975).

2. Approximating the generalized aerodynamic forces

Since a closed form solution for the aerodynamic forces on the plate does not exist, an approximation is sought. Recent work by Hall *et al.* (1995) and Romanowski and Dowell (1995) demonstrates that reduced order aerodynamic models can be constructed. The approach summarized here employs singular value decomposition (SVD) and was previously detailed by Frampton and Clark (1996). The basic concepts associated with the SVD are briefly reviewed here.

As indicated in Eq. (13), the force on the nth panel mode due to motion of the mth panel mode can be described in terms of two influence coefficients and two influence functions. The influence functions are convolved with the displacement and velocity of the panel response in generalized coordinates and, since no closed-form solutions for the influence functions exist, they must be obtained by numerical integration as outlined by Frampton *et al.* (1996a). The numerical solution of each aerodynamic influence function expressed in (13) can be cast in the form of a discrete-time impulse response function. A state-variable model of these aerodynamic impulse response functions can then be rendered by a technique previously outlined by Juang and Pappa (1985).

It so happens that a discrete-time state space realization of the form

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k), \tag{20}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k), \tag{21}$$

where the system ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$) has n states, p inputs, and q outputs, has a matrix-valued impulse response of

$$\mathbf{h}(k) = \begin{cases} \mathbf{D}, & \text{for } k=0, \\ \mathbf{C}\mathbf{A}^{k-1}\mathbf{B}, & \text{for } k=1,2,...,J+1 \\ \mathbf{0} & \text{for } k>J+1. \end{cases} \tag{22}$$

This formulation is unique to the discrete-time state-variable model. The block Hankel matrix can be formed from the matrix valued impulse response as

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(1) & \mathbf{h}(2) & \mathbf{h}(3) & \cdots & \mathbf{h}(J+1) \\ \mathbf{h}(2) & \mathbf{h}(3) & \mathbf{h}(4) & \cdots & \mathbf{0} \\ \mathbf{h}(3) & \mathbf{h}(4) & \mathbf{h}(5) & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}(J+1) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}, \tag{23}$$

where $\mathbf{H}$ is of dimension $q(J+1) \times p(J+1)$. The first step toward obtaining a state space realization of the system which is characterized by the matrix-valued impulse response is to perform a singular value decomposition of the block Hankel matrix. The decomposition has the form

$$\mathbf{H} = \bar{\mathbf{U}}\Sigma\bar{\mathbf{V}}^T = \mathbf{U}\mathbf{V}^T, \tag{24}$$

where $\mathbf{U}$ is a $q(J+1) \times (J+1)$ orthogonal matrix of the eigenvectors of $\mathbf{H}\mathbf{H}^T$, $\mathbf{V}$ is a $p(J+1) \times (J+1)$ orthogonal matrix of the eigenvectors of $\mathbf{H}^T\mathbf{H}$, and

$$\Sigma = \text{diag}[\sigma_1, \sigma_2, ..., \sigma_r, \epsilon_{r+1}, \epsilon_{r+2}, ..., \epsilon_{J+1}] \tag{25}$$

is a diagonal matrix of the singular values with

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \sigma_r \geqslant \epsilon_{r+1} \geqslant \epsilon_{r+2} \geqslant \cdots \geqslant \epsilon_{J+1} \tag{26}$$

and

$$\mathbf{U} = \bar{\mathbf{U}}\Sigma^{1/2}, \tag{27}$$

$$\mathbf{V}^T = \Sigma^{1/2}\bar{\mathbf{V}}^T. \tag{28}$$

If the matrix $\mathbf{H}$ has a rank of r, then the last $J-r$ singular values ($\epsilon_i : i = r+1, r+2, ..., J+1$) would be zero. If the singular values, ϵ_i, are not zero but are much less than the singular values, σ_i, then the matrix is very near to rank r. In this case the singular values, ϵ_i, represent computational ''noise'' in the impulse responses and degrees of freedom with relatively small contribution.

A reduced order realization ($\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$) of order r, which omits the excess degrees of freedom, can be obtained by partitioning the SVD as follows:

$$\mathbf{H} = \mathbf{U}_{IJ}\mathbf{V}_{IJ}, \tag{29}$$

where

$$\mathbf{U}_{IJ} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{1,2} \\ \mathbf{U}_{2,1} & \mathbf{U}_{2,2} \\ \vdots & \vdots \\ \mathbf{U}_{J,1} & \mathbf{U}_{J,2} \\ \mathbf{U}_{J+1,1} & \mathbf{U}_{J+1,2} \end{bmatrix} \tag{30}$$

and

$$\mathbf{V}_{IJ} = \begin{bmatrix} \mathbf{V}_{1,1}^T & \mathbf{V}_{2,1}^T & \cdots & \mathbf{V}_{J,1}^T & \mathbf{V}_{J+1,1}^T \\ \mathbf{V}_{1,2}^T & \mathbf{V}_{2,2}^T & \cdots & \mathbf{V}_{J,2}^T & \mathbf{V}_{J+1,2}^T \end{bmatrix}. \tag{31}$$

The block matrices $\mathbf{U}_{i,1}$; $i=1,2,...,J+1$ are $q \times r$, $\mathbf{U}_{i,2}$; $i=1,2,...,J+1$ are $q \times (J+r-1)$, $\mathbf{V}_{i,1}$; $i=1,2,...,J+1$ are $p \times r$, and $\mathbf{V}_{i,2}$; $i=1,2,...,J+1$ are $p \times (J+r-1)$.

A nonunique reduced order realization can be defined as follows:

$$\hat{\mathbf{A}} = \begin{bmatrix} \mathbf{U}_{1,1} \\ \mathbf{U}_{2,1} \\ \vdots \\ \mathbf{U}_{J,1} \end{bmatrix}^+ \begin{bmatrix} \mathbf{U}_{2,1} \\ \mathbf{U}_{3,1} \\ \vdots \\ \mathbf{U}_{J+1,1} \end{bmatrix} \tag{32}$$

(where the $[\cdot]^+$ notation indicates pseudoinverse)

$$\hat{\mathbf{B}} = \mathbf{V}_{1,1}^T, \tag{33}$$

$$\hat{\mathbf{C}} = \mathbf{U}_{1,1}, \tag{34}$$

and

$$\hat{\mathbf{D}} = \mathbf{h}(0). \tag{35}$$

Proofs of the above realization along with proofs of stability can be found in references by Kung (1978) and Juang and Pappa (1985). A description of the error bound of this realization can be found in the reference by Al-Saggaf and Franklin (1987).
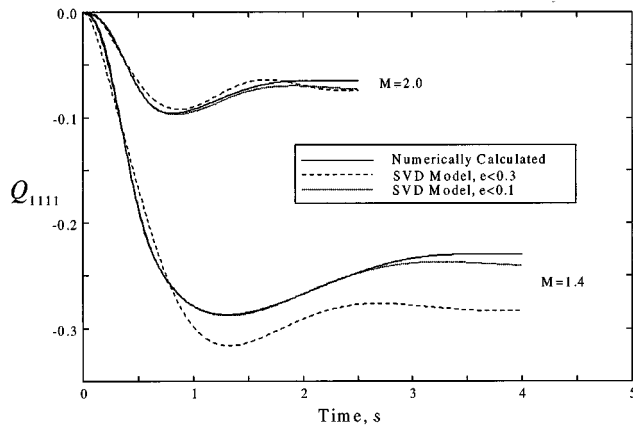
FIG. 4. Numeric and approximate generalized aerodynamic forces.

3. Application of SVD to an approximate aerodynamic model

The first step toward approximating the aerodynamic forces on a plate is to calculate the influence functions, $H_{mn}(t)$ and $I_{mn}(t)$. The integral of Eq. (19) can be performed analytically for a number of assumed panel modes; however, the integrals expressed in Eqs. (17) and (18) must be obtained through numerical integration (Frampton *et al.*, 1996a; Dowell, 1967, 1975). The numerical integration will provide a discrete-time representation of the aerodynamic influence functions. Next the matrix-valued impulse response for the aerodynamic system must be constructed. At this point, the block Hankel matrix of Eq. (23) can be constructed from the influence functions $H_{mn}(t)$ and $I_{mn}(t)$. A realization of the aerodynamic system can be then be formed and coupled to the plate system as depicted in Fig. 2.

A comparison of the generalized force on the (1,1) mode resulting from a step change in the (1,1) mode is shown in Fig. 4 versus nondimensional time, $s = t U_a / a$. Two reduced order models and the force calculated directly from Eq. (13) are shown. Note that the higher-order system, which resulted from a lower allowable error bound, follows the directly computed results well. A further increase in the system order would result in a response which is graphically indistinguishable from the directly computed result.

D. Radiation filter model

Radiation filters were originally introduced by Baumann *et al.* (1991) for active structural acoustic control such that the dynamic response of the system could be augmented to incorporate additional states (radiation modes) which provide a measure of sound power radiated. The objective of the work was to produce a cost functional in the LQR/LQG formulation which when minimized would guarantee optimal performance with respect to structural acoustic control as opposed to structural control. The augmentation of system dynamics in the formulation of a ''frequency shaped'' cost was originally proposed and detailed by Gupta (1980) and was extended to other exemplary applications. Thomas and Nelson (1995) also utilized the radiation filters for their feed-back control system realization, and the basic formulation for the filters can be found in their paper or that of Baumann *et al.* (1991).

For the infinite baffle, a frequency-dependent expression for the power radiated can be rendered by integrating an expression for the acoustic intensity over the hemisphere. Since the pressure can be obtained from Rayleigh's integral, the far-field acoustic intensity in the half-space is easily obtained. The form of the radiated power is

$$\Pi = \frac{1}{\pi} \int_0^\infty \mathbf{V}^H(jw) \mathbf{M}(jw) \mathbf{V}(jw) \, dw, \tag{36}$$

where Π is the total radiated power, $\mathbf{V}(j\omega)$ is matrix of the Fourier transform of modal velocities, and $\mathbf{M}(j\omega)$ represents the frequency-dependent interaction between structural modes required to compute the radiated power. The self- and mutual radiation efficiencies of the structural modes can be readily obtained from this matrix as outlined by Lomas and Hayek (1977) and more recently by Cunnefare (1991).

If the Laplace domain equivalent of the matrix $\mathbf{M}(j\omega)$ is expressed as $\mathbf{M}(s)$, then one can express the matrix in terms of a filter and its para-Hermitian conjugate:

$$\mathbf{M}(s) = \mathbf{G}^T(-s)\mathbf{G}(s), \tag{37}$$

where $\mathbf{G}(s)$ is the radiation filter as detailed by Baumann *et al.* (1991). This filter can be obtained through spectral factorization as outlined by Francis (1986), and a state-variable model for the filter can be used to augment the system dynamics. The time-averaged sound power radiated from the structure when driven by a white noise excitation can thus be computed from the output of these augmented states. For greater detail of the design of radiation filters, the reader is referred to the work by Baumann *et al.* (1991).

To demonstrate the accuracy of the spectral factorization, an exemplary plot of the radiation efficiency of the (1,1) structural mode is presented against the squared magnitude of the corresponding radiation filter in Fig. 5 (Wallace, 1972). As illustrated, the two curves are nearly indistinguishable. The remaining self- and mutual radiation efficiencies and corresponding radiation filters are of similar quality.

II. DESCRIPTION OF THE MODEL

The following parameters were used to construct the model used in this study: $a = b = 0.3$ m, $h_s = 0.0016$ m, $E_s = 7 \times 10^{10}$ Pa, $\rho_s = 2700$ kg/m^3, $\eta_s = 0.01$, $\rho_a = 0.42$ kg/m^3, and $a_a = 310$ m/s. The plate modes included in the model were the (1,1), (2,1), (3,1), and (4,1). As discussed earlier, the spanwise modes were neglected since it was assumed that the flow was uniform in the y direction of the panel, and most of the variation occurs in the chordwise direction. The material properties of the panel correspond to that of aluminum and the aerodynamic properties correspond to an altitude of 10 km.

Two disturbance inputs were considered in this study. The first provides a measure of the variation in transmission loss across the panel since all modes were excited uniformly and independently with a white noise process. This method of excitation ignores the spectral shaping associated with the
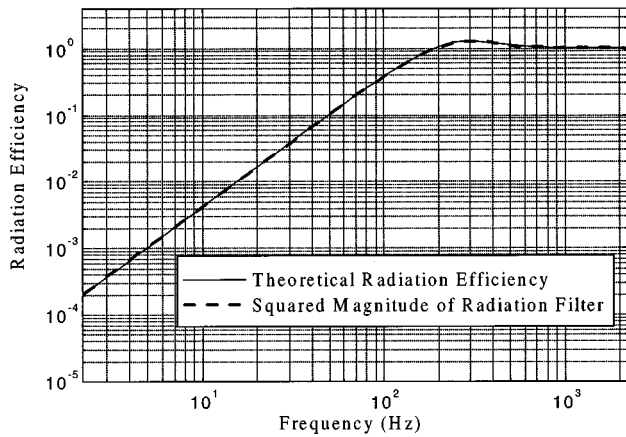
FIG. 5. Comparison of radiation efficiency of (1,1) mode computed numerically and with corresponding radiation filter.

FIG. 6. The rms sound power radiated by a panel subjected to subsonic flow and white noise excitation.

TBL model, and can be used to strictly determine the effects of aeroelasticity on the sound power radiated. The rms sound power is plotted as a function of frequency for various Mach numbers, and the integral of the sound power is presented numerically, which corresponds to the H_2-norm, and can be computed from the singular values of the system transfer function as follows (Boyd and Barratt, 1991):

$$\|H\|_2 = \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{i=1}^{n} \sigma_i(H(jw))^2 \, dw \right)^{1/2}. \quad (38)$$

Thus, once the system transfer function is represented in terms of state variables, it is reasonably trivial to compute the total rms sound power radiated when driven with independent white noise excitations. Results from this initial test case intentionally ignores flow-dependent scaling of the TBL noise source to isolate the panel dynamics as a function of flow speed.

The second set of results displays the effect that increasing Mach number has on the level of the TBL noise input. The TBL noise scales with the square of mean flow speed and thus must be considered in terms of predicting increases in sound power levels due to increases in the disturbance levels. This is an important test case to include in the results since one might naively assume that the sound power transmitted across the baffle decreases with increasing Mach number. In fact, the transmission loss increases due to aeroelasticity, but the disturbance noise source also increases and overwhelms the benefit associated with what has been frequently termed ''aerodynamic damping'' (Lyle and Dowell, 1994).

III. RESULTS

As discussed in the last section, two types of excitations in generalized coordinates were considered: white noise and white noise shaped by the TBL noise filters. The first case conveys the importance of including aeroelasticity in the dynamic model and the effect that aeroelasticity has on the sound power transmitted by the panel. The second case serves to demonstrate that the levels of the TBL noise source
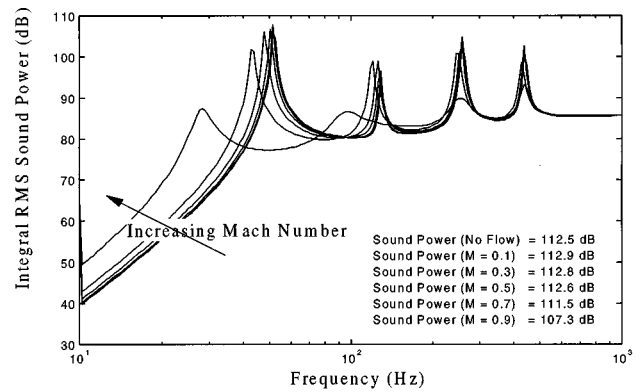
increase with increasing flow speed and thus any benefit gained through aeroelasticity near Mach 1 is offset by the increased level of the disturbance source.

A. Sound power radiated with white noise excitation

The first test case displays results from excitation of the panel in generalized coordinates with independent, spectrally white, noise sources. Results from this study are displayed in Fig. 6 for Mach numbers less than 1. As illustrated, the sound power radiated decreases with increasing Mach number, and at $M = 0.9$, the sound power radiated has decreased by approximately 7.4 dB for the example presented. This decrease in transmitted sound can be viewed in terms of an increase in transmission loss across the panel due to what has typically been termed aerodynamic damping (Lyle and Dowell, 1994) and is caused by an increase in the sound radiated to the convected fluid. While the level of damping certainly increases as the flow speed approaches $M = 1$, the stiffness of the panel modes is also modified, indicating an overall change in the structural impedance (Frampton and Clark, 1997). The resonant frequency associated with the fundamental mode of the panel moves significantly as a function of flow speed. Any feedback controller designed for active structural acoustic control must also include the effects of variations in panel dynamics as a function of flow speed. From this initial study, it appears that aeroelasticity actually serves a purpose in structural acoustic control when the flow speed is near $M = 1$.

The supersonic case is also considered, despite the fact that subsonic flow is currently of greater interest in the commercial industry. Again, a spectrally white noise source was used, and the Mach number was increased from $M = 1.2$ to $M = 2.0$. Results from this study are presented in Fig. 7. As illustrated, the resonant frequencies of the panel vary as a function of flow speed; however, in the supersonic case, there is no general trend for movement in the resonant frequencies as a function of increasing Mach number. In addition, one observes that there is relatively little change in the total sound power radiated by the panel in the supersonic case. It is clear from the results presented in Fig. 7 that if one were to design an active structural acoustic control system for supersonic conditions, variations in the plant dynamics
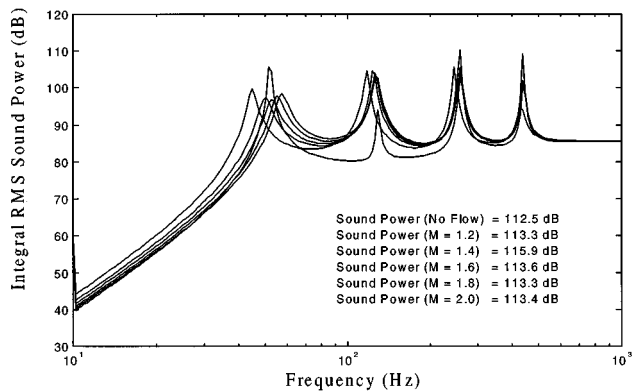
FIG. 7. The rms sound power radiated by a panel subjected to supersonic flow and white noise excitation.

FIG. 9. The rms sound power radiated by a panel subjected to supersonic flow and TBL noise excitation.

would need be included to ensure robustness of the controller over the desired flight conditions.

Note that for a sufficiently large supersonic Mach number flutter will occur (Dowell, 1975). Flutter is an inherent instability which is associated with the plate/fluid coupling. As the panel approaches the flutter boundary, one pair of poles will migrate into the right-half s plane and lead to instability. At this point, considering linear theory, the transmitted sound power will become infinite. However, in a real system the plate nonlinear effects will become prevalent and the system response will remain finite.

B. Sound power radiated with TBL noise excitation

As in the previous subsection, results for both subsonic and supersonic cases are presented. However, in this section, the sound power radiated as a function of the power spectrum associated with the TBL noise source is presented. Consider the power spectrum presented in Fig. 8. At low Mach number, the TBL noise source is very weak and thus the sound power radiated by the panel is limited to 38 dB. As the Mach number is increased, the sound power radiated also increases. Comparing these results to that presented in Fig. 6, one observes that although the transmission loss across the panel is increasing as a result of an increase in input impedance with increasing Mach number, the sound power transmitted also increases due to the fact that the TBL pressures

scale with the square of the mean velocity. As illustrated, the sound power transmitted increases by approximately 40 dB from $M = 0.1$ to $M = 0.9$, corresponding to an order of magnitude increase in flow speed. Thus, any benefit associated with increasing transmission loss as a function of flow speed is offset by the increase in the TBL pressures as a function of flow speed.

Results for supersonic flow speeds are illustrated in Fig. 9, and, as can be seen, the total sound power radiated continues to increase as a function of flow speed. For the results presented, a doubling of the flow speed in the supersonic region results in an approximate 10 dB increase in sound power transmitted. As in the subsonic case, this increase is dominated by the increase in TBL pressures with flow speed. Both the subsonic and supersonic test cases demonstrate that the TBL energy content is dominant at low frequency. However, a more significant ''roll-off'' in the high-frequency energy is present at supersonic speeds than at subsonic speeds. This ''roll-off'' occurs because the TBL power spectrum ''roll-off'' is greater over the frequency range of interest for the example presented as is observed upon comparing the TBL power spectrums in Fig. 3(a) and (b) for subsonic and supersonic flow, respectively.

IV. CONCLUSIONS

A method of modeling the aeroelastic structural acoustic coupling of a panel was presented in this work. One of the primary contributions of this effort rests in the realization of fluid, structure, and acoustic modeling in typical state-variable form. The formulation presented is amenable to modern and robust control system design and synthesis techniques.

Results from this study serve to emphasize the importance of including aeroelastic coupling in models used to predict performance of active structural acoustic control of aircraft panels. In both subsonic and supersonic flow conditions, the panel dynamics are significantly modified in the presence of external flow. As a result, these variations in plant dynamics must be considered in the design of robust control systems for active structural acoustic control.

At subsonic speeds, the aeroelastic coupling serves to increase the transmission loss across the panel with increas-

FIG. 8. The rms sound power radiated by a panel subjected to subsonic flow and TBL noise excitation.

ing Mach number due to what has typically been termed aerodynamic damping. However, the benefit of increasing flow speed associated with aeroelastic coupling is offset by the increase in the TBL noise source which scales with the square of the mean flow velocity. Thus, while the transmission loss is increased, the power spectrum of the disturbance also increases, and at a faster rate, leading to greater sound power transmission levels.

Future work will include the structural acoustic interaction of an enclosed sound field as opposed to radiation into a half-space. In addition, various robust control approaches will be investigated to determine performance bounds in the presence of ensuring robustness to variations in plant dynamics.

Al-Saggaf, U. M., and Franklin, G. F. (**1987**). ''An Error Bound for a Discrete Reduced Order Model of a Linear Multivariable System,'' IEEE Trans. Autom. Control. **AC-32**(9), 815–819.

Baumann, W. T., Saunders, W. R., and Robertshaw, H. H. (**1991**). ''Active suppression of acoustic radiation from impulsively excited structures,'' J. Acoust. Soc. Am. **90**, 3202–3208.

Blake, W. K. (**1986**). *Mechanics of Flow-induced Sound and Vibration. Volume II, Complex Flow-structure Interaction* (Academic, New York).

Boyd, S., and Barratt, C. (**1991**). *Linear Controller Design: Limits of Performance* (Prentice–Hall, Englewood Cliffs, NJ).

Bull, M. K. (**1996**). ''Wall-Pressure Fluctuations Beneath Turbulent Boundary Layers: Some Reflections on Forty Years of Research,'' J. Sound Vib. **190**(3), 299–315.

Corcos, G. M. (**1967**). ''The Resolution of Turbulent Pressures at the Wall of Boundary Layer,'' J. Sound Vib. **6**, 59–70.

Cunefare, K. A. (**1991**). ''The minimum multimodal radiation efficiency of baffled finite beams,'' J. Acoust. Soc. Am. **90**, 2521–2529.

Dowell, E. H. (**1967**). ''Generalized Aerodynamic Forces on a Flexible Plate Undergoing Transient Motion,'' Q. Appl. Math. **26**(3), 2267–2270.

Dowell, E. H. (**1975**). *Aeroelasticity of Plates and Shells* (Noordhoff, Groningen, The Netherlands).

Dugundji, J. (**1966**). ''Theoretical Considerations of Panel Flutter at High Supersonic Mach Numbers,'' AIAA J. **4**(7), 1257–1266.

Frampton, K. D., and Clark, R. L. (**1996**). ''State space modeling of aerodynamic forces on a plate using singular value decomposition,'' AIAA J. **34**(12), 2627–2630.

Frampton, K. D., and Clark, R. L. (**1997**). ''Sound Transmission through an Aeroelastic Plate into a Cavity,'' AIAA J. **35**(7), 1113–1118.

Frampton, K. D., Clark, R. L., and Dowell, E. H. (**1996a**). ''State-Space Modeling for Aeroelastic Panels Subject to Linearized Potential Flow Aerodynamic Loading,'' AIAA J. Aircraft **33**(4), 816–822.

Frampton, K. D., Clark, R. L., and Dowell, E. H. (**1996b**). ''Active Control of Panel Flutter with Linearized Potential Flow Aerodynamics,'' AIAA J. **33**(4), 768–774.

Francis, B. (**1986**). *A Course in H_∞ Control Theory* (Springer-Verlag, New York).

Gray, C. E., and Mei, C. (**1992**). ''Large Amplitude Finite Element Analysis of Composite Panels in Hypersonic Flow,'' *Proceedings of the AIAA Dynamics Specialist Conference*, Dallas, Texas.

Gupta, N. K. (**1980**). ''Frequency Shaped Cost Functionals: Extension of Linear-Quadratic-Gaussian Design Methods,'' J. Guid. Control. Dyn. **3**(6), 529–535.

Hajela, P. (**1991**). ''Application of Piezoelectric Elements in Supersonic Panel Flutter Suppression,'' *AIAA Aircraft Design System and Operations Meeting*, Baltimore, Maryland.

Hall, K. C., Florea, R., and Lanzkron, P. J. (**1995**). ''A Reduced Order Model of Unsteady Flows in Turbomachinery,'' J. Turbomach. **117**(3), 417–432.

Juang, J., and Pappa, R. S. (**1985**). ''An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction,'' J. Guid. Control. Dyn. **8**(9), 620–627.

Kung, S. (**1978**). ''A New Identification and Model Reduction Algorithm via Singular Value Decompositions,'' *IEEE 12th Asilomar Conference on Circuits, Systems and Computers*, pp. 705–714.

Lomas, N. S., and Hayek, S. I. (**1977**). ''Vibration and Acoustic Radiation of Elastically Supported Rectangular Plates,'' J. Sound Vib. **52**(1), 1–25.

Lyle, K. H., and Dowell, E. H. (**1994**). ''Acoustic Radiation Damping of Flat Rectangular Plates Subjected to Subsonic Flows. Parts I and II,'' J. Fluids Struct. **8**(7), 711–746.

Meirovitch, L. (**1967**). *Analytical Methods in Vibrations* (Macmillan, New York).

Romanowski, M. C., and Dowell, E. H. (**1995**). ''Aeroelastic Analysis of an Airfoil Using Eigenmode Based Reduced Order Unsteady Aerodynamics,'' AIAA Paper 95-1380.

Scott, R. C. (**1994**). ''Panel Flutter Suppression Using Adaptive Material Actuators,'' J. Aircraft **31**(1), 213–222.

Thomas, D. R., and Nelson, P. A. (**1995**). ''Feedback Control of Sound Radiation from a Plate Excited by a Turbulent Boundary Layer,'' J. Acoust. Soc. Am. **98**, 2651–2662.

Wallace, C. E. (**1972**). ''Radiation Resistance of a Rectangular Panel,'' J. Acoust. Soc. Am. **51**, 946–952.

Active control of a distributed-parameter structure using vortex power flow confinement

Nobuo Tanaka, Yoshihiro Kikushima, Masaharu Kuroda, and Neil J. Fergusson
Sound and Vibration Division, Mechanical Engineering Laboratory, Ministry of International Trade and Industry, 1-2 Namiki, Tsukuba Science City, Ibaraki 305, Japan

This paper considers the active power flow control of a distributed-parameter planar structure, with particular emphasis on a vortex power flow which has the potential to confine the vibrational power into a restricted area of the structure. Without letting the vibrational power, a cause of exciting structural modes, disperse into the structure, control of the structural response can be achieved. This paper begins by deriving the necessary condition for producing a vortex power flow in the vicinity of a disturbance point force at an arbitrary exciting frequency. Then, with a wave visualization system newly developed, it becomes possible to observe the wave propagation (the vortex power flow) taking place in the structure. In order to investigate the contribution of structural modes to the vortex power flow configuration, an energy stream function as well as a vorticity function is derived in a general form. By using these functions, the generation mechanism of a vortex power flow actively induced around a disturbance force location is quantitatively studied. © *1997 Acoustical Society of America.* [S0001-4966(97)00209-9]

PACS numbers: 43.40.Vn [PJR]

INTRODUCTION

In recent years, the active control of vibrational power flow[1-3] in distributed-parameter structures has become a topic of increased interest. When attempting to suppress the vibration of a distributed-parameter structure, previous active vibration control approaches, mostly based upon modal control[4,5] strategy, can no longer be applied due to an infinite number of vibration modes. As a matter of fact, the modal control-based approach, regardless of what control theory is employed: the classical control, modern control, or robust control, is applicable for cases where control of only a few modes are required. As far as modal control is concerned, there are as many sensors and actuators required for control as the number of vibration modes of a targeted structure. This paper presents a novel active control methodology, enabling control of a large number of structural modes.

Consider a lossless, distributed-parameter structure subject to an external force. The external force does not instantly excite the vibration modes. Rather, a certain procedure follows by which the vibration modes of a structure become completely excited. First, the external energy injected into a structure propagates through the structure as progressive waves, arriving at the structure boundaries. Then, in accordance with the characteristics of the boundary conditions of the structure, evanescent waves as well as reflected waves are produced. Now that there exist both incident and reflected waves in the structure, these waves interfere with each other, causing standing waves. If and only if the shapes of the standing waves are exactly the same as those of the natural modes will the modes be excited.

This discussion suggests two novel active vibration control schemes for a distributed-parameter structure, *progressive wave control*[6-9] and *standing wave control*.[6,10] In contrast to conventional active control whereby the vibration level at sensing points is attenuated, progressive wave control reduces vibration by eliminating its source, reflected waves, and hence the standing waves these reflected waves produce. In such a case, the entire infinity of vibration modes remain dormant. The progressive wave control methodology may thus be regarded as an implanting of the characteristics of an infinite structure, in which no vibration modes occur, onto a finite structure.

The second, or standing wave control technique, involves production of standing waves which never tune to the natural modes. In this instance, standing waves dominate the structural response, resulting in an absence of power flow in the lossless distributed-parameter structure.

This paper considers modified progressive wave control of a lossless structure. Unlike pure progressive wave control, once control is effected, both stationary and progressive waves, and hence nonzero power flow, persist, the power flow, which can be measured as vibration intensity,[11,12] serving as a performance index. Vibration intensity distribution patterns (or power flow patterns) appear in a lossless planar structure in three forms:[13] *a straight pattern, a figure-S pattern, and a vortex pattern*. In this paper we consider the latter type, which has the potential to confine external energy to a restricted area, or to divert power flow out of a specific region. After deriving a necessary condition for the presence of a vortex pattern in the vicinity of a disturbance, a wave visualization system is developed, enabling observation of the resulting power flow propagation.

In the process, an energy stream function, as well as a vorticity function, is derived in a general form. These functions permit quantitative discussion of the generation mechanism of the vortex power flow, as many as 100 vibration modes being deemed necessary to form a vortex power flow pattern.
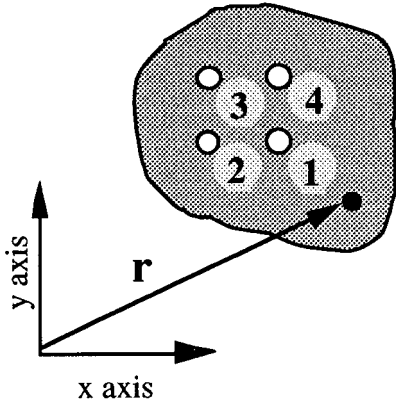
FIG. 1. Configuration of point forces.

I. INDUCING A VORTEX POWER FLOW

When two vibration modes are excited to similar levels whose resonance frequencies are very close to each other, a vortex power flow occurs.[14] The location where it occurs is determined by the interference of the two vibration modes. In other words, it is not possible to designate the vortex power flow location in a structure. It would be desirable if the power flow injected into a structure by a disturbance force could be confined to the vicinity of the disturbance location, without letting it disperse into the structure, thus providing sufficient control effect for suppressing the structural vibration. With this in mind, this article begins by detailing the method,[13] for inducing a vortex power flow around a disturbance point.

As shown in Fig. 1, let a set of four point forces be arranged in a quadrilateral fashion, with the bottom left force labeled as f_1, and the others numbered in a counterclockwise fashion. Consider a simply supported, thin, lossless, steel rectangular plate of lengths L_x and L_y in the x and y directions, respectively, and of thickness h. If a displacement $w(\mathbf{r})$ at the location $\mathbf{r}=(x,y)$ of a plate is written in terms of N_m vibration modes as

$$w(\mathbf{r})=\sum_{i=1}^{N_m} w_i \sin \alpha_i x \sin \beta_i y, \tag{1}$$

where the modal amplitude w_i and α_i and β_i are defined by

$$w_i=\frac{4}{\rho_s h L_x L_y(\omega_i^2-\omega^2)}\sum_{i=1}^{4} f_i \sin \alpha_i x \sin \beta_i y, \tag{2}$$

$$\alpha_i=\frac{m_i\pi}{L_x}, \tag{3}$$

$$\beta_i=\frac{n_i\pi}{L_y}, \tag{4}$$

ρ_s, ω_i, ω, m_i, n_i, and j being the material density, ith mode resonance frequency, excitation frequency, modal indices in the x and y direction, and the imaginary operator $\sqrt{-1}$, respectively.

The power flow at the point of application of the ith point force is given by

$$P_i=\tfrac{1}{2} \text{Re}[f_i \dot{w}^*(\mathbf{r}_i)], \quad i=1,2,3,4, \tag{5}$$

where *, ˙, and Re denote the complex conjugate, time derivative, and real part of the expression, respectively. By inserting Eqs. (1)–(4) into Eq. (5) and expanding Eq. (5), the local power flows are obtained as

$$P_1=-\frac{\omega}{2}\{s_{12} \text{Re}[jf_1f_2^*]+s_{13} \text{Re}[jf_1f_3^*]$$
$$+s_{14} \text{Re}[jf_1f_4^*]\},$$

$$P_2=-\frac{\omega}{2}\{s_{21} \text{Re}[jf_2f_1^*]+s_{23} \text{Re}[jf_2f_3^*]$$
$$+s_{24} \text{Re}[jf_2f_4^*]\}, \tag{6}$$

$$P_3=-\frac{\omega}{2}\{s_{31} \text{Re}[jf_3f_1^*]+s_{32} \text{Re}[jf_3f_2^*]$$
$$+s_{34} \text{Re}[jf_3f_4^*]\},$$

$$P_4=-\frac{\omega}{2}\{s_{41} \text{Re}[jf_4f_1^*]+s_{42} \text{Re}[jf_4f_2^*]$$
$$+s_{43} \text{Re}[jf_4f_3^*]\},$$

where

$$s_{lm}=\sum_{i=1}^{N_m} \frac{4}{\rho_s h L_x L_y(\omega_i^2-\omega^2)}$$
$$\times \sin \alpha_i x_l \sin \beta_i y_l \sin \alpha_i x_m \sin \beta_i y_m. \tag{7}$$

Furthermore, the terms in Eqs. (6) and (7) satisfy the following relationships:

$$\text{Re}[jf_lf_m^*]=-\text{Re}[jf_mf_l^*], \tag{8}$$

$$s_{lm}=s_{ml}. \tag{9}$$

Thus for the expressions in Eq. (6), Kirchhoff's law holds:

$$P_1+P_2+P_3+P_4=0. \tag{10}$$

Setting the power flow P_i equal to the power flow P_j at diametrically opposed locations i and j, the conditions necessary to produce a power flow vortex are

$$P_1=P_3, \tag{11}$$

$$P_2=P_4. \tag{12}$$

Using Eqs. (11) and (12), and Kirchhoff's law described in Eq. (10), the following relationships between the power flows are obtained:

$$P_1=-P_2, \tag{13}$$

$$P_3=-P_4. \tag{14}$$

Equations (11)–(14) show that a source and sink of the power flow arise in turn, and thus Eqs. (11) and (12) represent the necessary conditions for generation of a vortex power flow in the vicinity of the disturbance force.

Consider the control force necessary to produce the vortex power flow. Defining f_1 as the reference, which could also be regarded as a disturbance force, the other three forces used as control forces will be related to f_1 by the relationship

$$f_i=g_ie^{j\varphi_i}f_1, \quad i=2,3,4, \tag{15}$$
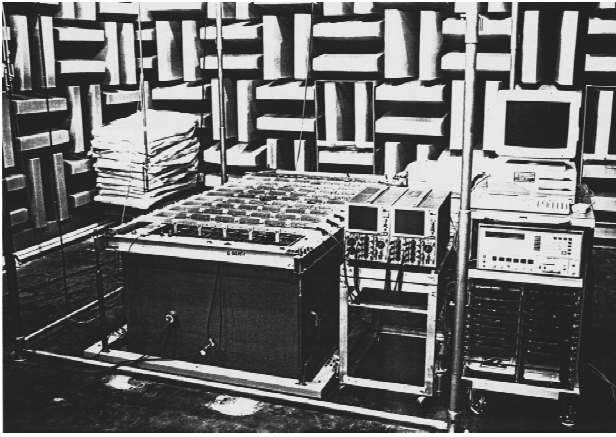
FIG. 2. Test rig for visualization of wave propagation.

FIG. 3. Vibration intensity distribution (experimental results).

where g_i and φ_i are the feedforward gain and phase of the ith control force, respectively. Note that there are four point forces acting on the plate, and that the number of equations representing the condition necessary to induce the vortex power flow is two. Thus the number of unknown parameters must be also two. By setting the phases, for instance, to $\varphi_2 = \pi/2$, $\varphi_3 = \pi$, and $\varphi_4 = 3\pi/2$, the force amplitudes are found to be

$$g_3 = \sqrt{\frac{s_{12}s_{14}}{s_{23}s_{34}}}\, g_1, \tag{16}$$

$$g_4 = \sqrt{\frac{s_{12}s_{23}}{s_{34}s_{14}}}\, g_2. \tag{17}$$

Thus, the above two expressions for the feedforward gains, g_3 and g_4, describe the necessary conditions for inducing a vortex power flow around the disturbance force location.

II. POWER FLOW CONFINEMENT

A. Wave visualization system

Figure 2 illustrates the experimental rig designed for inducing a vortex power flow. A steel rectangular panel with dimensions 88 cm×180 cm×9 mm is fixed on top of an enclosure constructed from 10-cm-thick concrete walls. The uppermost edges are each of a ''knife-blade'' type in order to simulate simply supported boundary conditions. The enclosed cavity is packed with glasswool to dampen acoustic resonances. Four electrodynamic shakers, one for primary excitation and the other three for control, are mounted on the bottom of the enclosure with rubber isolation pads, and attached to the plate using force transducers. Vibration intensity is measured using an intensity-meter and a pair of accelerometers, the latter being swapped for two microphones to measure acoustic intensity. With the aim of observing the wave propagation arising in the plate, a wave visualization system, consisting of a set of displacement sensors, a data logger, and a signal processor, is developed. As with a displacement sensor array, 55 noncontact displacement sensors (whose effective stroke is 0.5 mm and resolution is 0.5 μm) are placed above the panel with a grid size of 15 cm in both the x and y directions. Output signals of these displacement sensors are acquired by the data logger with a simultaneous sampling, and are transferred to the signal processor to interpolate those 55 data at each sampling in an effort to visualize two-dimensional propagation waveforms.
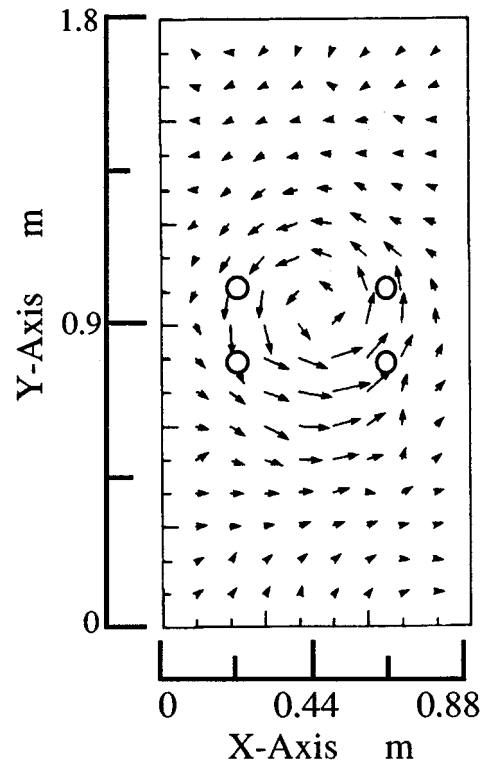
B. Measurement of a vortex power flow

For the purpose of producing a vortex power flow at the center of the plate, four electrodynamic shakers are installed at the center of the plate, 10 and 20.5 cm apart in the x and y directions, respectively. Using the necessary conditions described by Eqs. (16) and (17), the four excitation forces are determined: all gains of the forces are set to 1 N in amplitude of a sinusoidal wave; each phase difference between the adjacent forces is set to 90°; and the excitation frequency is set at 37 Hz, 99% of the (1,1) mode resonance frequency of the plate. Figure 3 shows the vibration intensity distribution obtained experimentally. As is to be expected, a vortex power flow is induced at the center of the plate, verifying the validity of the condition for inducing the vortex power flow as discussed in Sec. I.

Figure 4 depicts the time history of the vibrational responses for over half a period. As can be seen from the figure, a bump appearing at the center of the plate rotates counterclockwise, this movement corresponding to the vortex of vibration intensity shown in Fig. 3. Note that the (1, 2) modal pattern at $t = 0.4/f_{11}$ [f_{11} is the resonance frequency of the (1, 1) mode] and the (2, 1) modal pattern at $t = 0.2/f_{11}$ emerge in turn, demonstrating that these two modes contribute significantly to the generation of the vortex power flow. Using the wave visualization system, the envelope of the absolute amplitude of waves associated with each of these responses are depicted in three dimensions for the cases before and after control. For the case before control, where the
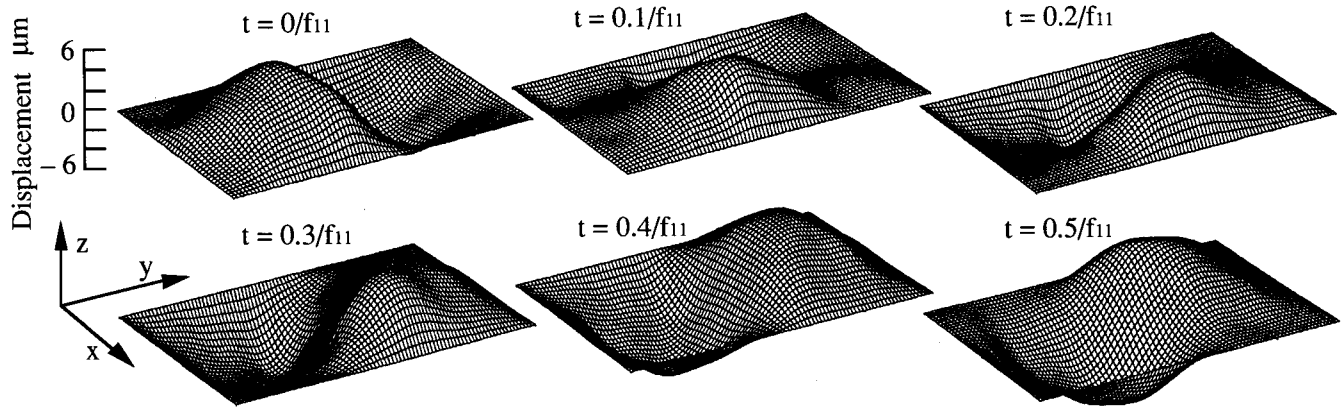
FIG. 4. Time history of propagation of waves around the center of the plate (experimental results).

only disturbance force acts at location 1 (see Fig. 1), the typical (1, 1) modal pattern is shown in the upper picture of Fig. 5. The middle picture in Fig. 5 illustrates the case after control, where the vortex power flow is generated at the center of the plate, while the lower one depicts the contour for this case. It is evident from the figure that the wave rotates around the center of the plate, with the displacement being zero at the center of the vortex. For the result shown in Fig. 5 the maximum displacement amplitude is 6.3 μm, 2.6% of the uncontrolled amplitude of 239 μm. Thus by confining the

FIG. 5. Envelope of absolute wave amplitude and contour for the cases before and after inducing vortex power flow at 35 Hz (experimental results).

power flow to a neighborhood of the disturbance force, a significant control effect can be obtained.

III. ENERGY FLOW FUNCTION AND VORTICITY FUNCTION

A. Derivation of energy stream function

It has been reported that a vortex power flow is induced[14] when two vibration modes whose resonance frequencies are very close to each other (like a degenerate mode) are excited simultaneously. In this case, the vortex location is given by the interference pattern of the vibration modes called a vortex block,[13] the two vibration modes determining the power flow configuration completely. Such a vortex is, as it were, a spontaneous vortex, while the vortex dealt with in this article is an actively induced one. Hence, it will be of interest to elucidate the generation mechanism of the actively induced vortex power flow, and to investigate how the structural modes are involved in the generation mechanism. To this end, this paper derives a generic form for the energy stream function[14,15] described by using N_m modes. With the energy stream function, it will become possible to see how the power flow path is constructed as well as how the vortex configuration evolves. With a view to clarifying the characteristics of the vortex, this paper further derives the vorticity function by way of which it will become possible to decide whether the vortex power flow actively generated is a real vortex or not.

For this, first and foremost, it is necessary to derive the vibration intensity[11,12] field. The vibration intensity vector $\mathbf{I}_v$ can be defined as

$$\mathbf{I}_v = [I_{vx}, I_{vy}]^T, \tag{18}$$

whose components I_{vx} and I_{vy} are given by

$$I_{vx} = \frac{1}{2} \text{Re} \left[Q_x(\mathbf{r}) \dot{w}^*(\mathbf{r}) + M_x(\mathbf{r}) \frac{\partial \dot{w}^*(\mathbf{r})}{\partial x} \right.$$
$$\left. + T_{xy}(\mathbf{r}) \frac{\partial \dot{w}^*(\mathbf{r})}{\partial y} \right], \tag{19}$$

$$I_{vy} = \frac{1}{2} \text{Re} \left[Q_y(\mathbf{r}) \dot{w}^*(\mathbf{r}) + M_y(\mathbf{r}) \frac{\partial \dot{w}^*(\mathbf{r})}{\partial y} \right.$$

$$\left. + T_{yx}(\mathbf{r}) \frac{\partial \dot{w}(\mathbf{r})}{\partial x} \right]. \tag{20}$$

Here, Q, M, and T are the shear force, bending moment, and torsional moment of the panel, respectively. The energy stream function ψ associated with $\mathbf{I}_v$ is now introduced according to

$$I_{vx} = \frac{\partial \psi}{\partial y}, \quad I_{vy} = -\frac{\partial \psi}{\partial x}, \tag{21}$$

or in terms of a differential form by

$$d\psi = -I_{vy} \, dx + I_{vx} \, dy. \tag{22}$$

By integrating Eq. (22) the energy stream function is obtained as

$$\psi = \int I_{vx} \, dy - \int I_{vy} \, dx. \tag{23}$$

Inserting Eqs. (19) and (20) into Eq. (23) and subsequently expanding leads to the energy stream function in terms of N_m vibration modes:

$$\psi = \frac{1}{2} \sum_{k=1}^{N_m} \sum_{l=1}^{N_m} \text{Re} \left[\psi_{kl}^x + \psi_{kl}^y \right]. \tag{24}$$

For the case when $k \neq l$

$$\psi_{kl}^x = \frac{1}{2} j\omega D w_k w_l^* \left[-(\alpha_k^2 + \beta_k^2)\alpha_k \cos \alpha_k x \right.$$

$$\times \sin \alpha_l x \left\{ \frac{\sin(\beta_k - \beta_l)y}{\beta_k - \beta_l} - \frac{\sin(\beta_k + \beta_l)y}{\beta_k + \beta_l} \right\}$$

$$+ (\alpha_k^2 + \nu \beta_k^2)\alpha \sin \alpha_k x \cos \alpha_l x$$

$$\times \left\{ \frac{\sin(\beta_k - \beta_l)y}{\beta_k - \beta_l} - \frac{\sin(\beta_k + \beta_l)y}{\beta_k + \beta_l} \right\}$$

$$- (1-\nu)\alpha_k \beta_k \beta_l \cos \alpha_k x \sin \alpha_x$$

$$\left. \times \left\{ \frac{\sin(\beta_k - \beta_l)y}{\beta_k - \beta_l} + \frac{\sin(\beta_k + \beta_l)y}{\beta_k + \beta_l} \right\} \right], \tag{25}$$

$$\psi_{kl}^y = \frac{1}{2} j\omega D w_k w_l^* \left[-(\alpha_k^2 + \beta_k^2)\beta_k \cos \beta_k y \right.$$

$$\times \sin \beta_l y \left\{ \frac{\sin(\alpha_k - \alpha_l)x}{\alpha_k - \alpha_l} - \frac{\sin(\alpha_k + \alpha_l)x}{\alpha_k + \alpha_l} \right\}$$

$$+ (\nu \alpha_k^2 + \beta_k^2)\beta_l \sin \beta_k y \cos \beta_l y$$

$$\times \left\{ \frac{\sin(\alpha_k - \alpha_l)x}{\alpha_k - \alpha_l} - \frac{\sin(\alpha_k + \alpha_l)x}{\alpha_k + \alpha_l} \right\}$$

$$- (1-\nu)\alpha_k \alpha_l \beta_k \cos \beta_k y \sin \beta_l y$$

$$\left. \times \left\{ \frac{\sin(\alpha_k - \alpha_l)x}{\alpha_k - \alpha_l} + \frac{\sin(\alpha_k + \alpha_l)x}{\alpha_k + \alpha_l} \right\} \right], \tag{26}$$

and for the case when $k = l$

$$\psi_{kl}^x = \frac{1}{2} j\omega D w_k w_l^* \left[-(\alpha_k^2 + \beta_k^2)\alpha_k \cos \alpha_k x \sin \alpha_l x \right.$$

$$\times \left\{ y - \frac{\sin(\beta_k + \beta_l)y}{\beta_k + \beta_l} \right\} + (\alpha_k^2 + \nu \beta_k^2)\alpha_l \sin \alpha_k x$$

$$\times \cos \alpha_l x \left\{ y - \frac{\sin(\beta_k + \beta_l)y}{\beta_k + \beta_l} \right\}$$

$$- (1-\nu)\alpha_k \beta_k \beta_l \cos \alpha_k x \sin \alpha_l x$$

$$\left. \times \left\{ y + \frac{\sin(\beta_k + \beta_l)y}{\beta_k + \beta_l} \right\} \right], \tag{27}$$

$$\psi_{kl}^y = \frac{1}{2} j\omega D w_k w_l^* \left[-(\alpha_k^2 + \beta_k^2)\beta_k \cos \beta_k y \sin \beta_l y \right.$$

$$\times \left\{ x - \frac{\sin(\alpha_k + \alpha_l)x}{\alpha_k + \alpha_l} \right\} + (\nu \alpha_k^2 + \beta_k^2)\beta_l$$

$$\times \sin \beta_k y \cos \beta_l y \left\{ x - \frac{\sin(\alpha_k + \alpha_l)x}{\alpha_k + \alpha_l} \right\}$$

$$- (1-\nu)\alpha_k \alpha_l \beta_k \cos \beta_k y \sin \beta_l y$$

$$\left. \times \left\{ x + \frac{\sin(\alpha_k + \alpha_l)x}{\alpha_k + \alpha_l} \right\} \right], \tag{28}$$

where D and v denote the bending stiffness and a Poisson's ratio of the panel, respectively.

B. Derivation of the vorticity function

Given the energy stream function, a vorticity function can be introduced

$$\zeta = -\frac{1}{2} \left(\frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi}{\partial y^2} \right). \tag{29}$$

Inserting Eqs. (19), (20), and (23) into Eq. (29) and further expanding Eq. (29), the vorticity function ζ in terms of N_m vibration modes becomes

$$\zeta = \frac{1}{2} \sum_{k=1}^{N_m} \sum_{l=1}^{N_m} \text{Re} \left[(\zeta_{kl}^x + \zeta_{kl}^y) \right], \tag{30}$$

where

$$\zeta_{kl}^x = \tfrac{1}{2} j\omega D w_k w_l^* [-(\alpha_k^2 + \beta_k^2)\alpha_k \cos \alpha_k x \sin \alpha_l x$$

$$\times \{-(\beta_k - \beta_l) \sin(\beta_k - \beta_l)y + (\beta_k + \beta_l) \sin(\beta_k$$

$$+ \beta_l)y\} + (\alpha_k^2 + \nu \beta_k^2)\alpha_l \sin \alpha_k x \cos \alpha_l x \{-(\beta_k$$

$$- \beta_l) \sin(\beta_k - \beta_l)y + (\beta_k + \beta_l) \sin(\beta_k + \beta_l)y\}$$

$$+ (1-\nu)\alpha_k \beta_k \beta_l \cos \alpha_k x \sin \alpha_l x \{(\beta_k - \beta_l)$$

$$\times \sin(\beta_k - \beta_l)y + (\beta_k + \beta_l) \sin(\beta_k + \beta_l)y\}], \tag{31}$$

1652   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Tanaka *et al.*: Active control of structure   1652
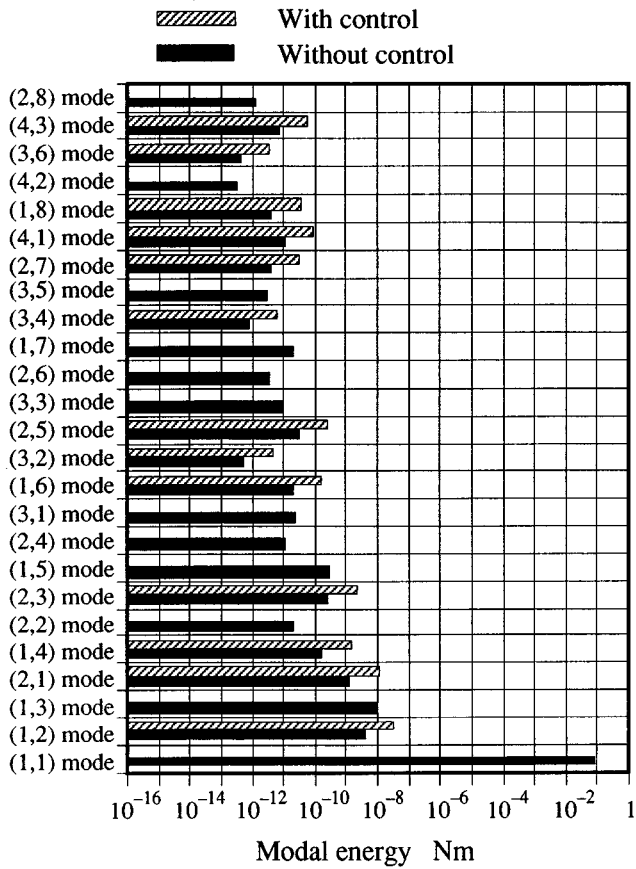
FIG. 6. Modal energy versus vibration modes.
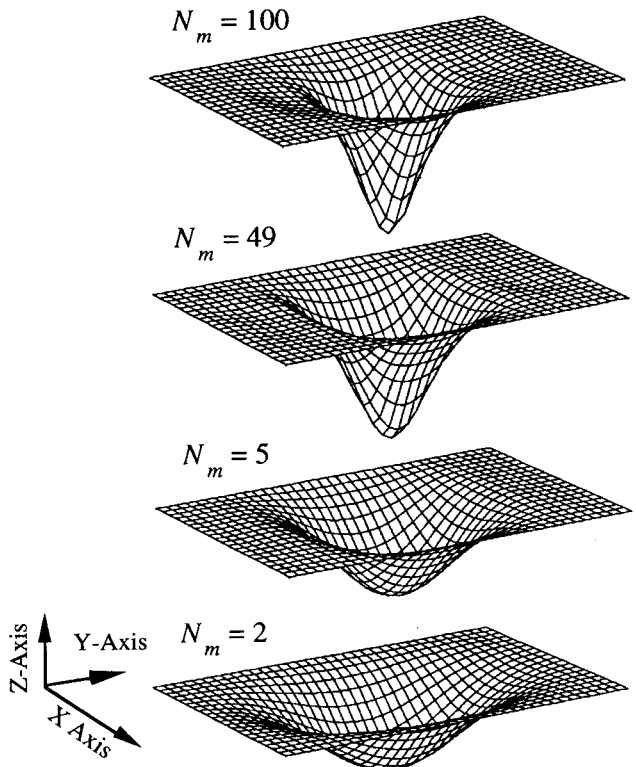
FIG. 7. Evolution of energy stream functions with the number of vibration modes.
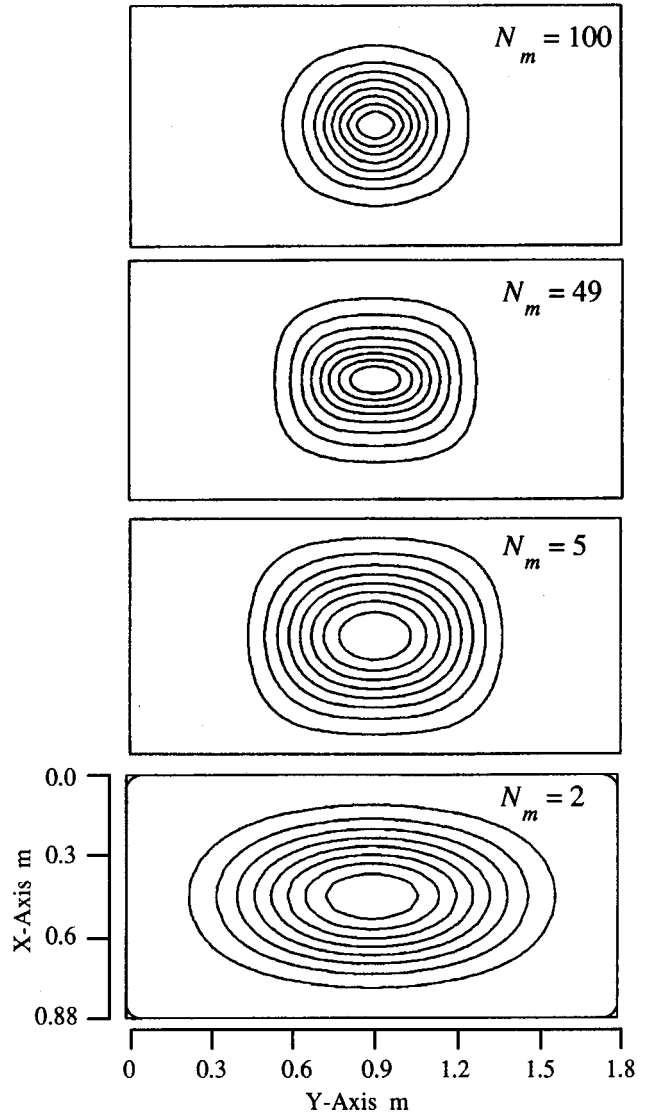
FIG. 8. Contour of energy stream functions evolving with a change of the number of vibration modes.

$$\zeta_{kl}^{y} = \tfrac{1}{2} j\omega D w_k w_l^* [-(\alpha_k^2 + \beta_k^2)\beta_k \cos \beta_k y \sin \beta_l y$$
$$\times \{ -(\alpha_k - \alpha_l)\sin(\alpha_k - \alpha_l)x + (\alpha_k + \alpha_l)$$
$$\times \sin(\alpha_k + \alpha_l)x \} + (\nu\alpha_k^2 + \beta_k^2)\beta_l \sin \beta_k y \cos \beta_l y$$
$$\times \{ -(\alpha_k - \alpha_l)\sin(\alpha_k - \alpha_l)x + (\alpha_k + \alpha_l)\sin(\alpha_k$$
$$+ \alpha_l)x \} + (1 - \nu)\alpha_k \alpha_l \beta_k \cos \beta_k y \sin \beta_l y \{ (\alpha_k$$
$$- \alpha_l)\sin(\alpha_k - \alpha_l)x + (\alpha_k + \alpha_l)\sin(\alpha_k + \alpha_l)x \}].$$

$$(32)$$

C. Modal energy

Figure 6 illustrates the modal energy[14] of each vibration mode of the plate with and without the vortex power flow generated at the center of the panel. The vertical axis of the figure indicates modal indices arranged from the lowest resonance frequency of the panel to the highest, while the horizontal axis shows the associated modal energy of the panel. As the exciting frequency of the primary force is close
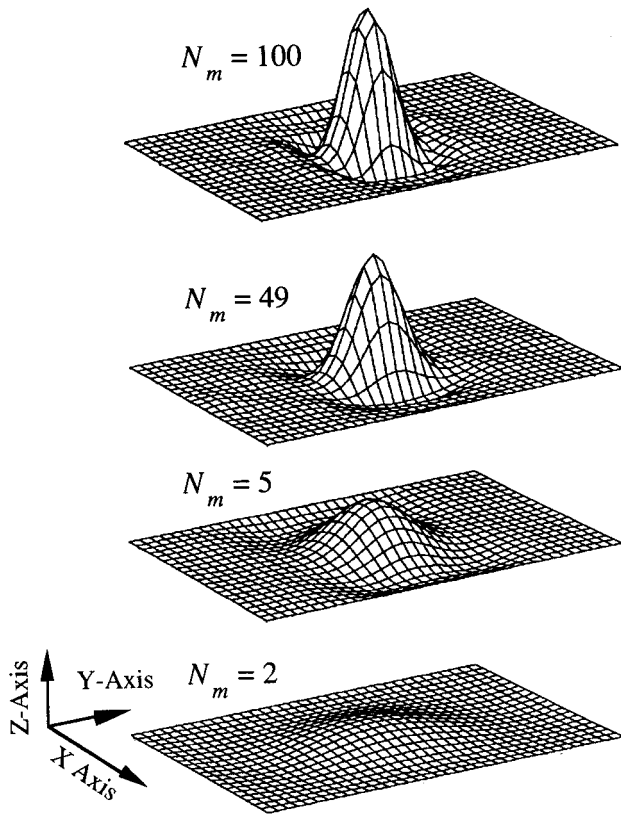
FIG. 9. Evolution of vorticity with the number of vibration modes.

FIG. 10. Contour of vorticity functions evolving with a change of the number of vibration modes (solid line, positive values; dotted line, negative values).

enough to the (1, 1) mode resonance frequency (99% of its frequency), the modal energy of the (1, 1) mode is conspicuous. For the case where a vortex power flow is induced at the center of the plate, however, the modal energy of the (1, 1) mode is perfectly suppressed.

Observe that when generating the vortex power flow, the modal energies of the (1, 3), (2, 3), (1, 5) modes are also completely suppressed. In contrast, the modal energies of the (1, 2), (2, 1), (1, 4) modes increase. From Fig. 6 it can be seen that the energy is highest for the (1, 2) modal, followed by the (2, 1) mode. Recall that the (1, 2) and (2, 1) modal patterns emerged in turn in the time history of the structural response as depicted in Fig. 4. The modal energies of these two vibration modes are the largest and the second largest contributor to the total energy of the plate with the vortex power flow at the center of the plate; these two modal patterns are thus visible in the figure. It should also be noted that, in contrast to the vortex power flow induced by the previously discussed case[13] of two close frequencies, the two frequencies [55.7 Hz for the (1, 2) mode and 120.7 Hz for the (2, 1) mode] associated with this actively induced vortex power flow differ by 70 Hz. By utilizing the vibration modes selected from the greater contributor to the total energy, the energy stream function can be evaluated. Figures 7 and 8 show how the energy stream function as well as its corresponding contour evolves with the number of vibration modes used from the greatest contributor to the total energy. As is clear from these figures, using only two vibration modes, the (1, 2) and (2, 1) modes, the foundation of the vortex power flow is already fabricated at the center of the

panel. The shape of the contour is, however, elliptic with a longer axis in the y direction. By increasing the number of contributing modes, the contour changes its form from an ellipse to a circle which is close to the experimentally obtained vibration intensity distribution pattern illustrated in Fig. 3. Gavric et al.[16] reported that at least 1000 vibration modes are needed for convergence of the vibration intensity distribution. Regarding the actively induced vortex power flow, however, 100 vibration modes are needed to form a precise power flow pattern. It should be noted that the modal energy of the (1, 2) mode, the greatest contributor to the total energy, is 1000 times as large as that of the tenth one, the (2, 7) mode; 3.44×10^{-8} Nm for the (1, 2) mode, but only 3.41×10^{-11} Nm for the (2, 7) mode. In other words, the tenth vibration mode contributes merely 0.1% of the (1, 2) modal energy to the total energy. A mode such as the 100th mode contributes even less. However, even such small contributors to the total energy are needed to form the precise vortex power flow configuration.
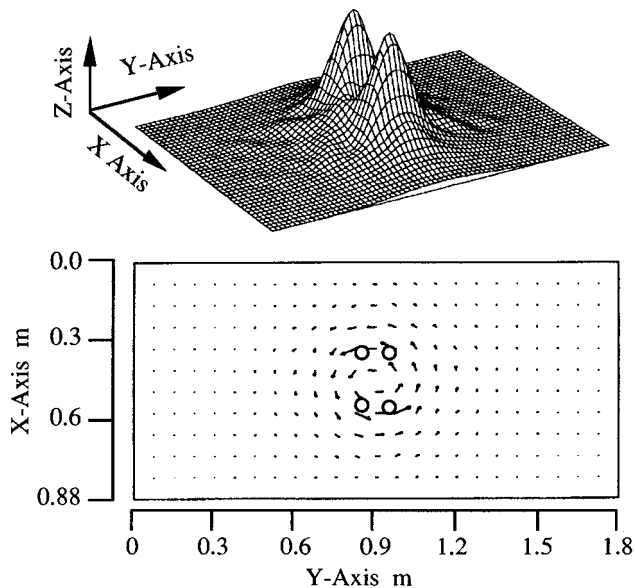
FIG. 11. (a) The 3-D figure of vibration intensity with a vortex at the center of the plate and (b) the vibration intensity distribution.

IV. DISCUSSION ON THE VORTEX POWER FLOW

The vorticity function can be likewise obtained by using the vibration modes in order of their contribution rate to the total energy. Figure 9 shows a three-dimensional picture of the vorticity function, demonstrating its evolution with the number of vibration modes; the associated contours are depicted in Fig. 10. It can be seen from these figures that only two vibration modes are sufficient to construct an approximate form of the vorticity function at the center of the panel. Analogous to the energy stream function, the contour of the vorticity function, using two vibration modes, has an elliptical form, altering its configuration from an ellipse to a circle as the number of the modes increases.

Recall the degenerate modal case[13] where subvortices exist inside a major vortex. If this is the case, the vorticity of each subvortex is nonzero, thereby the subvortex is a real vortex. The vorticity of the major vortex, however, turns out to be zero; hence it is not so much a real vortex as a derivative of the subvortices. On the contrary, the actively induced vortex generated at the center of the plate shown in Fig. 10 has vorticity, and hence it is a real vortex.

Figure 11 shows the numerical analysis of a vibration intensity distribution and its associated three-dimensional picture for the case when the vortex is produced at the center of the plate. In observing Figs. 3 and 11, the agreement between the experimental and theoretical results is seen to be in general good, except around the excitation points due to the influence of a nearfield. During an experiment, as the vibration intensity is measured by using a ''pair-of-accelerometers technique,'' the precision of the vibration intensity near the excitation points cannot be guaranteed. Furthermore, as seen from Fig. 11, the intensity value becomes zero at the center of the vortex. As the distance from the vortex center increases, the intensity level increases and we obtain a maximum around the exciting points; as the distance further increases, it then decreases.

With this phenomenon, the vortex characteristics can be investigated from the viewpoint of fluid dynamics; the region where the vibration intensity level increases as the distance from the vortex center increases is characterized by a forced vortex, while the vortex outside the region is characterized by a free vortex. Thus, the actively induced vortex shown here is called a combination vortex of forced and free vortices, as observed in fluid dynamics.

V. CONCLUSION

A vortex power flow has been successfully generated experimentally at the center of a simply supported rectangular panel. As the result, the power flow causing the excitation of the vibration modes is confined in the vicinity of a disturbance point in this case, and so the vibration level is significantly suppressed. The maximum amplitude of the panel without power flow control is 239 μm, and is suppressed to only 6.3 μm by the power flow control. With the wave visualization system developed here, observation of the wave propagation in a planar structure becomes possible. Furthermore, an energy stream function as well as a vorticity function is derived in a general form. With these functions, the generation mechanism of the actively induced vortex power flow is quantitatively discussed. It is also revealed that even the structural modes which contribute little to the total energy of a structure are needed to construct a power flow path. For the vortex case actively induced at the center of the panel, 100 vibration modes are deemed essential to form a circular vortex power flow path which agrees with the experimental power flow. Finally, the characteristics of a vortex power flow are investigated, showing that the vortex is characterized by a combination vortex consisting of a forced vortex and a free vortex.

[1] X. Pan and C. H. Hansen, ''Active control of power flow along a semi-infinite plate,'' J. Sound Vib. **184**(4), 585–610 (1995).
[2] J. Q. Pan and C. H. Hansen, ''Active control of vibratory power flow from a vibrating rigid body to a flexible panel through two active isolators,'' J. Acoust. Soc. Am. **93**, 1947–1953 (1993).
[3] N. Tanaka, Y. Kikushima, and M. Kuroda, ''Active wave control of a flexible beam (On the power flow control),'' J. Jpn. Soc. Mech. Eng. **58** (552), 2578–2585 (1992).
[4] L. Meirovitch and H. Baruh, ''Control of self-adjoint distributed parameter systems,'' J. Guid. Control. Dyn. **51**, 60–66 (1982).
[5] L. Meirovitch and S. Thangitham, ''The implementation of modal filters for control of structures,'' J. Guid. Control. Dyn. **8**(6), 706–716 (1985).
[6] N. Tanaka and Y. Kikushima, ''Active wave control of a flexible beam (Fundamental characteristics of an active sink system and its verification),'' JSME Int. J. **35**(2), 236–244 (1992).
[7] D. R. Vaughan, ''Application of distributed parameter concepts to dynamic analysis and control of bending vibrations,'' ASME, J. Basic Eng. **157**, 166 (1968).
[8] G. M. Procopio and J. E. Hubbard, Jr., ''Active damping of a Bernoulli–Euler beam via end point impedance control using distributed parameter techniques,'' ASME Design Technology Conference, 35–46 (1987).
[9] A. H. von Flotow and B. Schafer, ''Wave absorbing controllers for a flexible beam,'' J. Guidance **9**(6), 673–680 (1986).
[10] N. Tanaka, Y. Kikushima, M. Kuroda, and S. D. Snyder, ''Structure-borne intensity characteristics of an actively controlled thin plate,'' in 4th International Congress on Intensity Techniques, SENLIS (1993), pp. 273–280.
[11] D. U. Noiseux, ''Measurement of power flow in uniform beams and

plates,'' J. Acoust. Soc. Am. **47**, 238–247 (1970).

[12] G. Pavic, ''Measurement of structure borne wave intensity, Part I: Formulation of the methods,'' J. Sound Vib. **49**(2), 221–230 (1976).

[13] N. Tanaka, S. D. Snyder, Y. Kikushima, and M. Kuroda, ''Vortex structural power flow in a thin plate and the influence on the acoustic field,'' J. Acoust. Soc. Am. **96**, 1563–1574 (1994).

[14] N. Tanaka, S. D. Snyder, and C. H. Hansen, ''Vorticity of the vibrational intensity field in an actively controlled thin plate,'' J. Acoust. Soc. Am. **99**, 942–953 (1996).

[15] R. V. Waterhouse and D. Feit, ''Equal-energy streamlines,'' J. Acoust. Soc. Am. **80**, 681–684 (1986).

[16] L. Gavric and G. Pavic, ''A finite element method for computation of structural intensity by the normal mode approach,'' J. Sound Vib. **164**(1), 29–43 (1993).

Performance of noise indices in air-conditioned landscaped office buildings

S. K. Tang

Department of Building Services Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Noise measurements and a questionnaire survey were carried out in air-conditioned landscaped offices in the present study in an attempt to find out what existing noise index gives the best correlation with the auditory sensation feeling of office workers. Results from statistical analyses show that the *Equivalent Sound Pressure Level* is the best among the 14 commonly used noise indices for this purpose. It is also found that the less commonly used *Zwicker's loudness level* performs better than the commonly adopted *Noise Criterion curves* and *Noise Rating curves* in predicting auditory sensation of office workers in air-conditioned landscaped offices. The performance of the 14 noise indices in predicting auditory comfort are compared. © *1997 Acoustical Society of America.* [S0001-4966(97)05508-2]

PACS numbers: 43.50.Ba [GAD]

INTRODUCTION

One of the most controversial issues in building up a good indoor environment is the setting up of a suitable noise criterion. Noise affects human beings both physiologically and psychologically as discussed in Kryter.[1] A poor acoustical environment will lead to excessive mental stress, loss of concentration, etc., and, for commercial offices, these symptoms mean loss of productivity. The importance of noise control in buildings is reiterated by Lord and Wilson[2] in a discussion on the environmental design of buildings.

Noise inside a building comes from human activities, building services, and outdoors. The contribution of outdoor noise depends very much on the sound transmission loss of the building facade. Examples of research and surveys on this topic include Craik and Thancanamootoo[3] and Elmallawany.[4] However, most people today are working in air-conditioned office buildings whose facades have relatively higher sound transmission loss so that sound transmission from the outdoors to the indoors is insignificant when compared to those generated by the building services, such as the air-conditioning system, and human activities.[5]

A good acoustical environment implies good control of noise to keep its level and frequency content within the acceptable range for human beings. This issue deals with human feeling, which is very subjective in nature. However, some objective parameters are required for the design and assessment of an indoor acoustical environment. Many different noise indices have been developed to suit this purpose, among which the well-known equivalent sound pressure level $L_{Aeq, T = 5 \text{ min.}}$, the *noise Criterion curves (NC)* of Beranek,[6] and *noise rating curves (NR)* of Kosten and van Os[7] are commonly used nowadays. These indices are easy to interpret and measure. However, the search for better indices never ceases. Studies on the noise inside air-conditioned office buildings have been rigorous for many decades. Keighley[8] gave a discussion on the determination of acceptability criteria for office noise while Hay and Kemp[9] proposed an office noise index based on percentile noise level,[1]

which is not commonly used today. There are still many other noise indices that may be able to predict human comfort effectively. This list includes the *room criteria (RC)* of Blaizer,[10] *loudness levels* of Stevens[11] and Zwicker,[12] *balanced noise criteria (NCB)* of Beranek,[13] etc. Details of this list will be given later. A particular review on the criterion for controlling indoor air-conditioning noise can be found in Kingsbury.[14] Because of the existence of so many noise indices, it is important to test how well they can represent human auditory rating for future use in the setting up of suitable noise index for building environmental noise control. It is assumed at this stage that their performance may depend on the type of environment. The present study focuses on air-conditioned landscaped offices.

In the present study, sound level and spectral measurements of noise in one-third octave bands were carried out in parallel with a questionnaire survey in 26 air-conditioned landscaped offices during their normal hours of operation. Therefore, the noise data in this study include not only the contribution from the air-conditioning system, but also those from human activities and office machines. The performance of various noise indices in correlating with the auditory rating of the office workers is discussed. It is hoped that the most effective index and a clue for future development of noise index for office buildings can be found.

I. NOISE MEASUREMENT AND SURVEY

Noise levels at different locations in the 26 air-conditioned landscaped offices were measured using a Brüel & Kjær 2236C precision sound level meter while their spectral content in one-third octave bands were obtained by Brüel & Kjær 2144 real-time frequency analyzer with Brüel & Kjær 4145 condenser microphone. The questionnaire asked the office workers to rate their feeling towards the acoustical environment on an unbiased linear seven-point scale at the time the measurements were being done. The present adopted scale is similar to that used by Tang *et al.*[5] and is shown in the Appendix. The rating obtained from the scale,

TABLE I. Abbreviations of noise indices.

Noise indices	Abbreviation	Unit
Equivalent sound pressure level	$L_{Aeq,T=5\,min.}$	dB
Percentile level for 10% exceedence	$L_{A10,T=5\,min.}$	dB
Percentile level for 90% exceedence	$L_{A90,T=5\,min.}$	dB
Noise climate	$L_{A10,T=5\,min.} - L_{A90,T=5\,min.}$	dB
Noise criterion	NC	Nil
Preferred noise criterion	PNC	Nil
Balanced noise criterion	NCB	Nil
Noise rating	NR	Nil
Room criterion	RC	Nil
Midfrequency average level	L_{MF}	dB
Stevens' loudness	LD_s	sone
Loudness level (Stevens)	LL_s	phon
Zwicker's loudness	LD_z	sone
Loudness level (Zwicker)	LL_z	phon

which ranges from -3 to $+3$, represents human auditory subjective rating. A rating of -3 represents ''Very Noisy'' feeling while ratings of 0 and $+3$ denote, respectively, the ''Neutral'' and ''Very Quiet'' feelings. One can observe that it is much simpler than that used by Beranek[6] so as to minimize disturbance to the office workers. The ratings so obtained represent their purely primitive but subjective feelings toward their acoustical environments as a whole. It should be noted that there is no guideline for the office workers to judge their local acoustical environments so that they may base their judgment on loudness or noisiness of office noise. This rating scale is also much different from the biased scale adopted by Ko *et al.*[15] in the study of air-conditioning nuisance; that of Ko *et al.*[15] mainly deals with complaints.

Each noise measurement in the present study lasted for 5 mins. A total of 1187 questionnaire responses were obtained. Each of them was accompanied by physical noise measurements.

II. NOISE INDICES

All the noise indices presented in the foregoing discussions, except $L_{Aeq,T=5\,min.}$ and those related to the percentile noise levels $L_{A10,T=5\,min.}$ and $L_{A90,T=5\,min.}$ were calculated from the one-third octave band noise spectra measured by the Brüel & Kjær 2144 real-time frequency analyzer. The noise indices involved in the present study are $L_{Aeq,T=5\,min.}$, $L_{A10,T=5\,min.}$, $L_{A90,T=5\,min.}$, the noise climate $L_{A10,T=5\,min.} - L_{A90,T=5\,min.}$, *noise criterion*,[6] *noise rating*,[7] *balanced noise criterion*,[13] *preferred noise criterion*,[16] *midfrequency average level* (L_{MF}),[17] *loudness*, and *loudness levels* of Stevens[11] and Zwicker.[12] The first four indices are A-weighted. However, one should bear in mind that these indices may not be mutually exclusive, and this list is by no mean exhaustive. There are still some other less common indices like the *low-frequency noise rating* proposed by Broner and Leventhall[18] which is specially designed to cater for low-frequency noise environment. *Articulation index*[1] is also not presented in the present study as it is believed that this index deals with speech intelligibility and is less likely to be associated with nuisance directly. Though L_{MF} can also deal with speech intelligibility,[1] it is the arithmetic mean of the octave band levels at 500 Hz, 1 kHz, and 2 kHz which

cover the most sensitive frequency range of human ear, and thus directly represents certain strength of a noise. It should also be noted from the results of Tachibana *et al.*[19] that the arithmetic average of the levels in the bands from 63 Hz to 4 kHz may give slightly better correlation with human feeling than *Zwicker's loudness level*. However, such difference in the performance of these two indices is not expected to be substantial, especially under the very subjective nature of human feeling. Thus, the index purposed by Tachibana *et al.*[19] will not be discussed in the present study.

The noise indices, except the noise level data $L_{Aeq,T=5\,min.}$, $L_{A10,T=5\,min.}$, $L_{A90,T=5\,min.}$, and the noise climate which were recorded *in situ* using the sound level meter, were computed from the spectra using a self-developed computer program. The indices NC, NCB, PNC, RC, and NR were determined by the ''tangency method.'' The FORTRAN subroutine for the calculation of *Zwicker's loudness* given by Paulus and Zwicker[20] was adopted. The method for Stevens' loudness calculation simply followed that given in ISO 532.[21] For ease of reference in the foregoing discussions, the noise indices are abbreviated as shown in Table I. Suffices *s* and *z* represent Stevens and Zwicker systems of loudness estimation, respectively. The loudness level LL in phons and loudness LD in sones, no matter which loudness system is concerned, are related by the expression[21]

$$LD = 2^{(LL-40)/10}.$$

III. RESULTS AND DISCUSSIONS

Before going into the details of the discussion on the performance of noise indices, some typical octave band noise spectra obtained in the offices during normal operation hours are shown in Fig. 1. It can be observed that the noise spectra in the surveyed offices vary significantly. However, most of them show hissy nature. This is due to office activities as suggested by Fig. 2. In general, over 85% of the noise spectra obtained in the present field measurements show hissy characteristics (not shown here). Therefore, though it is well known that NCB and the composite RC rating provide a number as well as a letter showing the frequency content of a noise; the latter is ignored in the foregoing discussion.
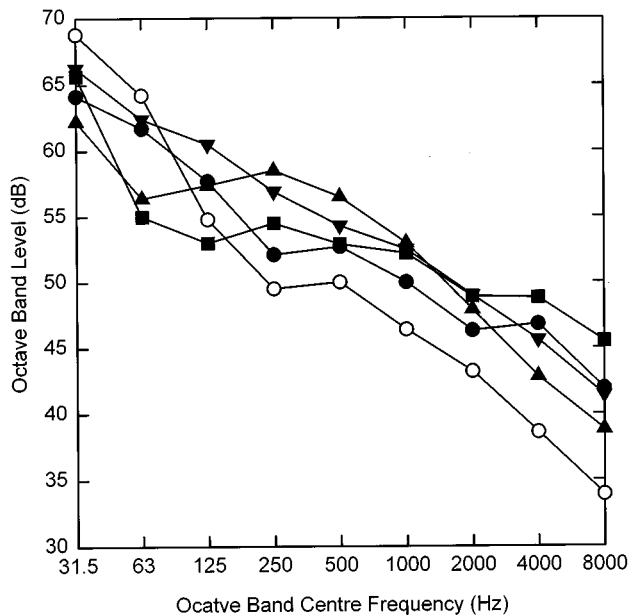
FIG. 1. Typical office noise spectra during normal operation hours. ○: Neutral spectrum; ●, ■, ▲, and ▼: hissy spectra.

The performance of noise indices is investigated through linear regression analysis. Good performance of a noise index in the present study refers to the case in which its correlation with the human auditory subjective rating is good and linear relationship with human auditory subjective rating can be established. Correlation coefficient is used for the assessment of noise index performance. Hypothesis testing procedure with the F-test at 95% confidence level illustrated in Bethea and Rhinehart[22] is also adopted to test whether the linear relationships between noise indices and the auditory sensation rating are valid. The null hypothesis H_0 is that the slope of the regression line is zero (that is, no linear relationship exists) and will be accepted if $F \leqslant F_{1,n-2,0.95}$, where n is the sample size. Only linear regression is considered in this study as it is the simplest form of analysis and visual inspection of the plots do not indicate clearly any nonlinear
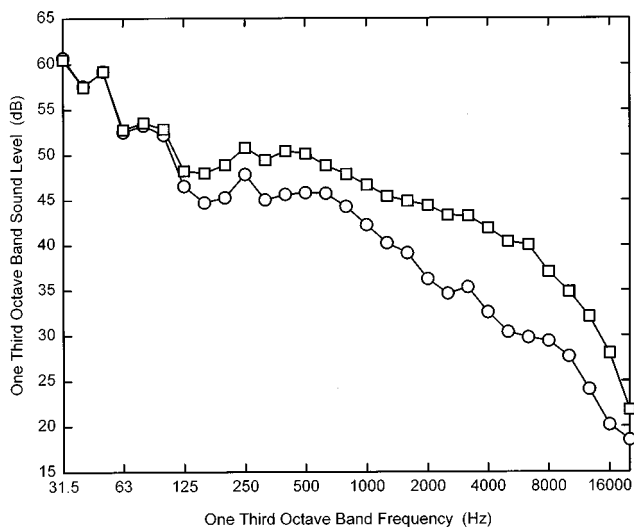
FIG. 2. Change in noise spectral content due to office activities. □: Noise spectrum during office hours; ○: noise spectrum outside office hours.

TABLE II. Interval for human auditory subjective rating averaging.

Noise index	Interval of averaging	Range
Equivalent sound pressure level	±0.5 dB	41–70 dB
Percentile level for 10% exceedence	±0.5 dB	42–75 dB
Percentile level for 90% exceedence	±0.5 dB	35–59 dB
Noise climate	±0.5 dB	2–18 dB
Noise criterion	Not applicable	NC 39–59
Preferred noise criterion	Not applicable	PNC 35–65
Balanced noise criterion	Not applicable	NCB 34–64
Noise rating	Not applicable	NR 33–70
Room criterion	Not applicable	RC 40–67
Midfrequency average level	±0.5 dB	31–74 dB
Stevens' loudness	±0.5 phon	15–130 sone
Loudness level (Stevens)	±0.5 sone	80–110 phon
Zwicker's loudness	±0.5 sone	5–30 sone
Loudness level (Zwicker)	±0.5 phon	62–89 phon

relationship. Due to the very subjective nature of human feeling, the method for rejecting outlying observations developed by Thompson[23] is adopted. Mean sensation ratings obtained from a sample of size less than five are ignored in the analysis. One index is better than another index when its R^2 and the associated ratio $F/F_{1,n-2,0.95}$ are larger than those of the latter.

Before discussing the regression results, it should be emphasized that two noises having the same $L_{\text{Aeq},T=5\,\text{min.}}$ will not necessarily have the same NC value because of the difference in the meanings of the two indices. This phenomenon applies to all other noise indices. Also owing to the subjective nature of human feeling, different office workers may give different ratings to the same acoustical environment. In the foregoing analysis, the human auditory sensation ratings presented are arithmetic averages within short intervals of noise indices. For example, the rating at $L_{\text{Aeq},T=5\,\text{min.}} = P$ dB is the arithmetic average of all votes obtained within the interval $P - 0.5 \leqslant L_{\text{Aeq},T=5\,\text{min.}} < P + 0.5$ (a ±0.5 dB interval). As there is no fractional value for NC, NCB, NR, PNC, and RC, the ratings presented with them are the straightforward means. Table II summarizes the interval of averaging and the range for each index. The noise indices discussed in the present study can basically be divided into three categories. The first category includes $L_{\text{Aeq},T=5\,\text{min.}}$, percentile levels, and the noise climate whose estimation does not require a knowledge of the noise spectral contents. The second one consists of *loudness, loudness levels*, and the *midfrequency average levels*, which are calculated from the spectral information of noise and are continuous indices. The last one refers to the group of NC, NCB, and etc. which have discrete nature and are, like those in the second category, estimated from the noise spectra. Though the emphasis of these indices is on achieving acceptable speech intelligibility and a nonannoyance shape of the noise spectrum for that speech intelligibility level,[17] it is still worthwhile to see whether they can be used to predict the very primitive but subjective feeling of office workers toward their acoustical environments.

Figure 3(a) shows the correlation between the average human auditory sensation rating S and $L_{\text{Aeq},T=5\,\text{min.}}$. The regression line is also shown in the figure. Though a certain
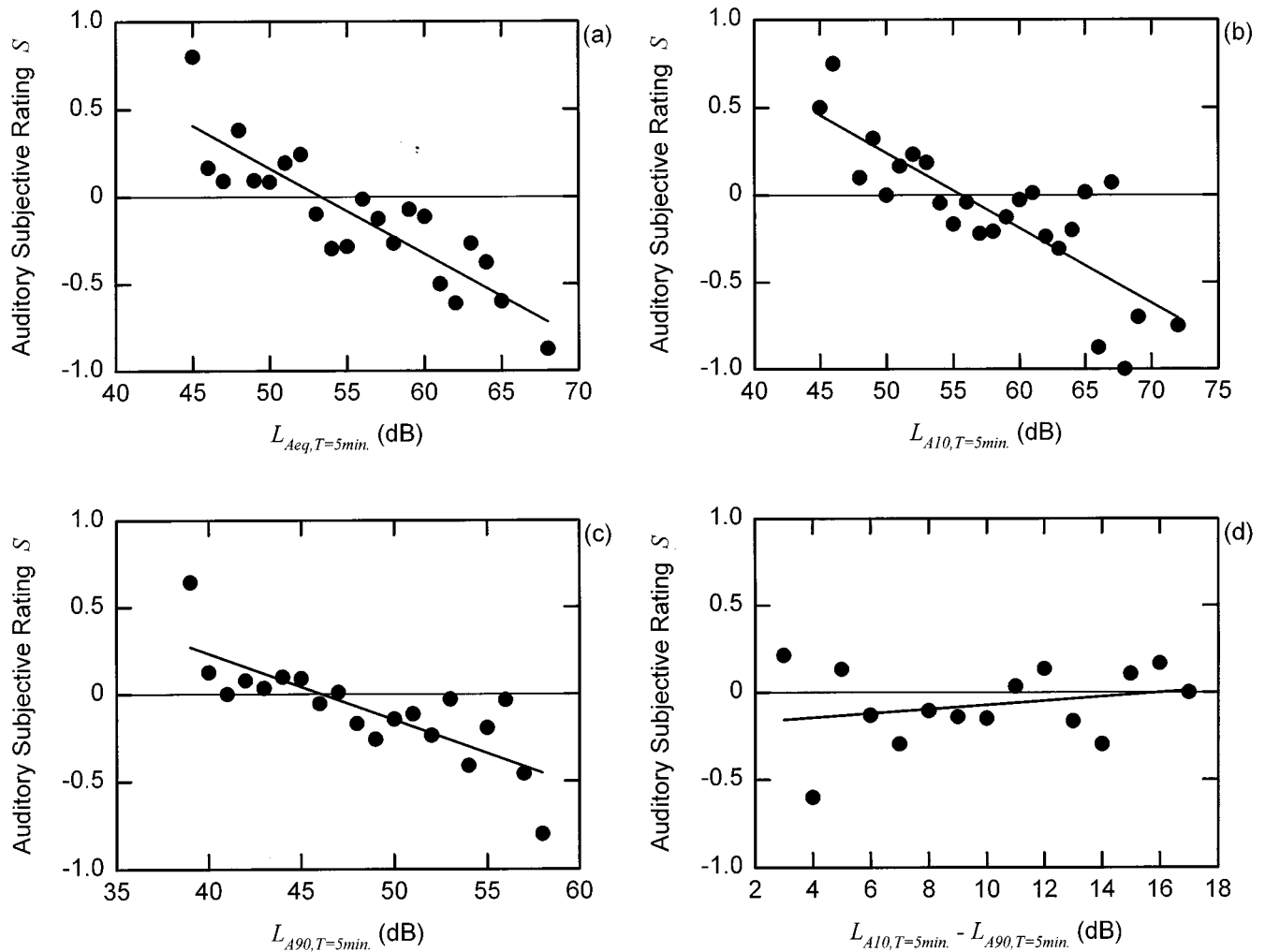
FIG. 3. Correlations between noise levels and human auditory subjective ratings. (a) $L_{Aeq, T=5 \text{ min.}}$; (b) $L_{A10, T=5 \text{ min.}}$; (c) $L_{A90, T=5 \text{ min.}}$; and (d) $L_{A10, T=5 \text{ min.}} - L_{A90, T=5 \text{ min.}}$. $\bullet$: Experimental data; ———: regression line.

degree of scatter is observed, the correlation coefficient R^2 is 0.766, which is relatively high in social research. Also, $F = 65.5 > F_{1,20,0.95} (=4.35)$ (Ref. 22) in the hypothesis testing and thus the null hypothesis is rejected, showing that linear relationship between S and $L_{Aeq,T=5 \text{ min.}}$ exists and the regression line should represent such relationship. The regression line predicts neutral auditory sensation at $L_{Aeq,T=5 \text{ min.}} \approx 52$ dB, which is slightly higher than the recommendation of 50 dB in BS 8233.[24] However, this difference is minor in an environmental noise study. The present predicted neutral auditory sensation does agree with the upper limit of L_{Aeq} suggested in Reynold[25] for large offices. Similar correlation phenomena are observed for $L_{A10,T=5 \text{ min.}}$ and $L_{A90,T=5 \text{ min.}}$ as shown in Fig. 3(b) and (c), respectively. Though the correlation coefficients between S and these two quantities are lower than that between S and $L_{Aeq,T=5 \text{ min.}}$ the results of the F-test are $F = 44.1$ for $L_{A10,T=5 \text{ min.}}$ and 30.1 for $L_{A90,T=5 \text{ min.}}$ showing again that the regression lines do represent the existence of linear relationships. However, the noise climate $L_{A10,T=5 \text{ min.}} - L_{A90, T=5 \text{ min.}}$ does not relate to human sensation S as shown in Fig. 3(d) ($R^2 = 0.06, F = 0.83 < F_{1,13,0.95} = 4.67$). This suggests that those noise indices which are associated with the noise climate may not be

appropriate for use in air-conditioned landscaped offices. Examples of such noise indices are office noise index proposed by Hay and Kemp[9] and the noise pollution level of Robinson.[26] Results of the present statistical analysis are tabulated in Table III. In Table III, n represents the number of averaging intervals. This number varies from index to index because of the range and average interval of each index (see Table II). It should also be noted that there are missing values in the noise index distribution so that n is, in general, less than the value obtained by dividing the range by the averaging interval. Figure 4 illustrates the frequency distributions of $L_{Aeq,T=5 \text{ min.}}$, $L_{A10,T=5 \text{ min.}}$, $L_{A90,T=5 \text{ min.}}$, and $L_{A10,T=5 \text{ min.}} - L_{A90T=5 \text{ min.}}$. Allowing for data scatter, they appear to be quite symmetrical (skewness $s \approx 0.4$). However, that of the *noise climate* $L_{A10,T=5 \text{ min.}} - L_{A90,T=5 \text{ min.}}$ is relatively more skewed to the low value side.

Indices in the second category are, in general, not very well correlated with human auditory sensation as shown in Table III. Only *Zwicker's loudness level* LL_z[12] can be considered as a useful index for application in air-conditioned landscaped offices. Others are either inappropriate or very marginal. The correlation between LL_z and human subjective rating S is shown in Fig. 5(a) ($R^2 = 0.65$). Unlike the fre-
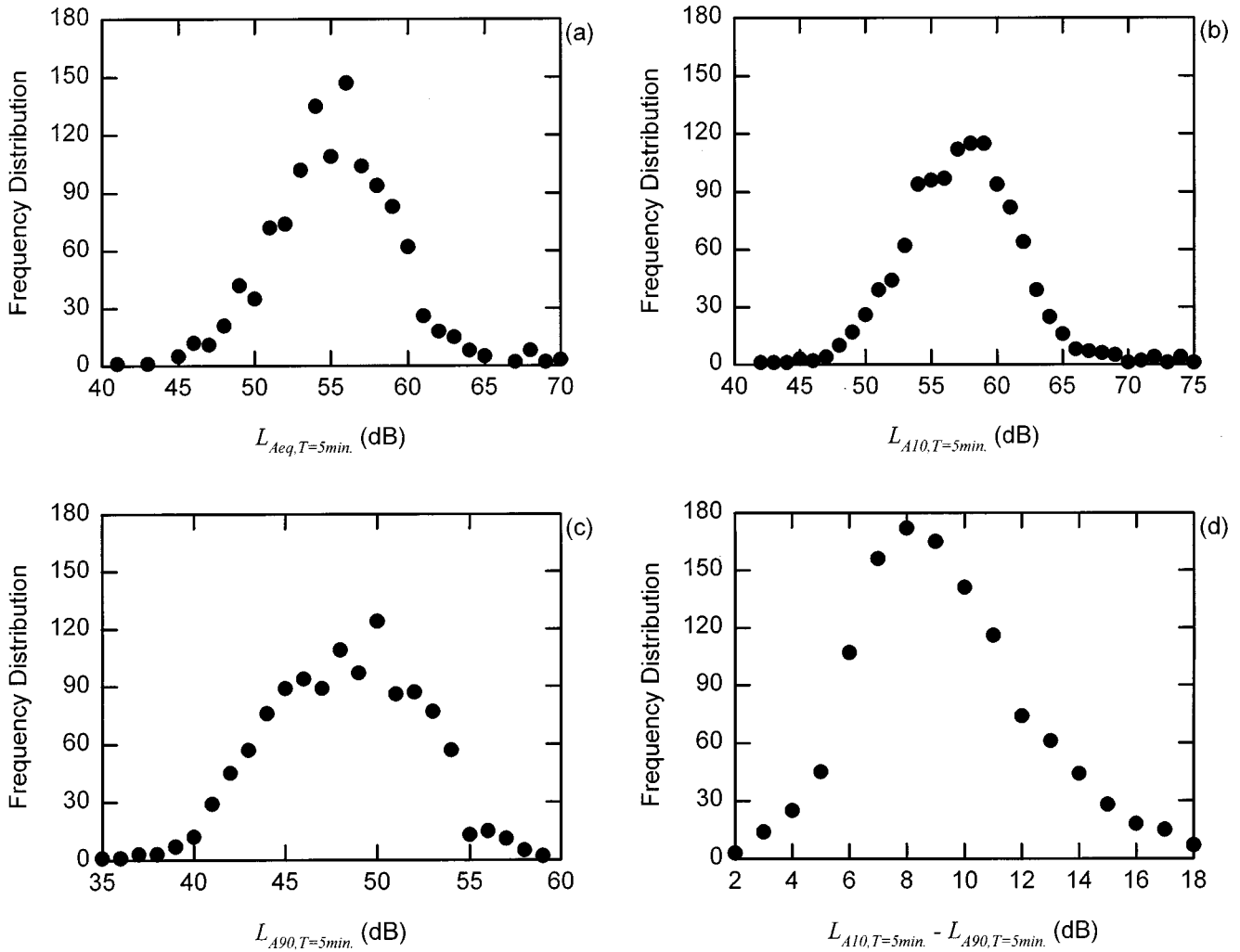
FIG. 4. Frequency distribution of noise levels. (a) $L_{\text{Aeq}, T=5\,\text{min.}}$; (b) $L_{\text{A10}, T=5\,\text{min.}}$; (c) $L_{\text{A90}, T=5\,\text{min.}}$; (d) $L_{\text{A10}, T=5\,\text{min.}} - L_{\text{A90}, T=5\,\text{min.}}$.

quency distributions of $L_{\text{Aeq},T=5\,\text{min.}}$, $L_{\text{A}10,T=5\,\text{min.}}$, and $L_{\text{A}90,T=5\,\text{min.}}$, those for noise indices within this category tend to have relatively higher kurtosis (for instance, 4.17 for LL_z, 5.41 for LL_s, and 2.66 for L_{MF}), showing that their

occurrence is concentrated in relatively more narrow ranges [Fig. 5(b)]. As only LL_z appears to be related to human auditory feeling, the frequency distributions of other noise indices within this category are not presented. Figure 5(a) sug-

TABLE III. Statistical result summary.

Noise indices	Skewness (s)	Kurtosis (k)	Number of averaging intervals n	Correlation coefficient (R^2)	Value of f-test (F)	$F_{1,n-2,0.95}$	H_0
$L_{\text{Aeq},T=5\,\text{min.}}$	0.532	1.940	22	0.766	65.45	4.35	Rejected
$L_{\text{A}10,T=5\,\text{min.}}$	0.457	1.700	25	0.657	44.09	4.24	Rejected
$L_{\text{A}90,T=5\,\text{min.}}$	0.368	1.233	20	0.626	30.07	4.41	Rejected
Noise climate	0.705	0.837	15	0.060	0.836	4.67	Accepted
NC	0.096	0.630	22	0.319	10.90	4.35	Marginal
PNC	−0.132	0.340	20	0.254	6.142	4.41	Marginal
NCB	−0.252	0.775	22	0.351	10.80	4.35	Marginal
NR	0.955	4.447	23	0.362	11.93	4.32	Marginal
RC	0.927	4.941	29	0.124	3.808	4.21	Accepted
L_{MF}	0.579	2.658	20	0.306	7.942	4.41	Marginal
LD_s	4.187	34.33	45	0.174	9.088	4.06	Marginal
LL_s	0.046	5.412	19	0.195	4.115	4.45	Accepted
LD_z	4.441	36.48	17	0.386	9.409	4.54	Marginal
LL_z	0.475	4.169	22	0.650	37.22	4.35	Rejected

FIG. 5. (a) Correlation between *Zwicker's loudness level* and auditory subjective rating. (b) Frequency distribution of *Zwicker's loudness level*. ●: Experimental data; ————: regression line.

FIG. 6. Correlations between NC, NCB, and NR with auditory sensation vote. ————: Regression line for NC; – – –: regression line for NCB; –·–: regression line for NR; ○: experimental data for NC; □: experimental data for NCB; △: experimental data for NR.

dicted by the marginally acceptable correlation in Fig. 6 which is higher than that of 45 shown in Reynold.[25]

Statistical results in Table III suggest that the equivalent sound pressure level $L_{Aeq,T=5\,min.}$ gives the best estimation of the auditory subjective rating of the office workers among all the noise indices discussed in the present study at 95% confidence level. The second best index studied in the present investigation is LL_z. Results also illustrate that the performances of discrete indices NC, NCB, NR, PNC, and RC in matching subjective human auditory feeling are, in general, not satisfactory.

IV. CONCLUSIONS

In the present investigation, noise measurements and a questionnaire survey were carried out in more than 20 air-conditioned landscaped offices. The questionnaire was used for the collection of subjective ratings of the office workers towards their acoustical environment. Physical noise measurement results included the equivalent sound pressure level, the percentile levels, and the one-third octave noise spectra from which various noise indices were calculated. Performance of noise indices in correlating with the subjective human auditory subjective rating is discussed.

The noise indices considered in the present study are $L_{Aeq,T=5\,min.}$, $L_{A10,T=5\,min.}$, $L_{A90,T=5\,min.}$, *noise climate, noise criterion curves (NC), noise rating curves (NR), the balanced noise criterion, preferred noise criterion, room criterion, midfrequency average level, loudness,* and *loudness level*. Statistical results show that among these indices, $L_{Aeq,T=5\,min.}$ correlates best with the auditory sensation. The second best is *Zwicker's loudness level*, which seems to have been ignored in the setting up of building noise control design guidelines nowadays. Though NC and NR have been widely used in the control of air-conditioning noise spectra and indoor environmental design, their correlations with human auditory sensation are not satisfactory. Their use in predicting human auditory comfort is only marginally accept-

gests that office workers will regard the acoustical environment to be "noisy" when $LL_z > 74$ phons. This *Zwicker's loudness level* has not been specified, at least to the knowledge of the author, in international standards or recommendations for indoor building environmental design and noise control. The good correlation between it and human auditory subjective rating manifests its importance in dealing with human hearing comfort in air-conditioned landscaped offices.

Table III also suggests that the noise indices in the third category, which have discrete nature, do not perform very well in the prediction of human auditory sensation in air-conditioned landscaped offices. The relatively better indices are NC, NCB, and NR. For RC, the correlation is essentially not valid. Also though F for PNC$= 6.14 > F_{1,n-2,0.95} = 4.41$, the difference is too small for a reliable rejection of the null hypothesis. Figure 6 illustrates the degree of correlation between NC, NCB, NR, and human auditory subjective rating. Though these correlations are only marginally acceptable, the present results suggest that office workers, as a whole, will feel noisy when NCB>45. This agrees with the recommendation of Beranek.[27] The neutral NC value is 47 as pre-

able. All other noise indices are found inappropriate for use in noise control for air-conditioned landscaped offices, especially the noise climate and room criterion which actually do not correlate significantly with the human sensation.

As the present investigation was performed in Hong Kong where most office workers are Chinese, it remains to find out whether similar results can be obtained in other cities.

ACKNOWLEDGMENTS

APPENDIX: SCALE FOR AUDITORY SUBJECTIVE RATINGS COLLECTION

The seven-point scale for the collection of the auditory subjective ratings of office workers in the present questionnaire survey is as shown below:

```
     -3      -2      -1       0      +1      +2      +3
      |       |       |       |       |       |       |
Very Noisy                 Neutral              Very Quiet
```
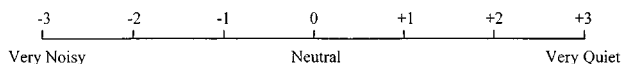
[1] K. D. Kryter, *The Effects of Noise on Man* (Academic, New York, 1985).

[2] E. A. Lord and C. B. Wilson, ''Description and prediction in the environmental design of buildings,'' Bldg. Envir. **17**, 293–300 (1982).

[3] R. J. M. Craik and A. Thancanamootoo, ''The importance of in-plane waves in sound transmission through buildings,'' Appl. Acoust. **37**, 85–109 (1992).

[4] A. Elmallawany, ''Field investigations of the sound insulation in school buildings,'' Bldg. Envir. **18**, 85–89 (1983).

[5] S. K. Tang, J. Burnett, and C. M. Poon, ''A survey on the aural environment in air conditioned open-plan offices,'' Bldg. Serv. Eng. Res. Tech. **17**, 97–100 (1996).

[6] L. L. Beranek, ''Criteria for office quieting based on questionnaire rating studies,'' J. Acoust. Soc. Am. **28**, 833–852 (1956).

[7] C. W. Kosten and G. J. van Os, ''Community reaction criteria to external noise,'' in National Physical Lab. Symp. No. 12 (Her Majesty's Stationary Office, London, 1962), pp. 373–387.

[8] E. C. Keighley, ''The determination of acceptability criteria for office noise,'' J. Sound Vib. **4**, 73–87 (1966).

[9] B. Hay and M. F. Kemp, ''Measurements of noise in air conditioned, landscaped offices,'' J. Sound Vib. **23**, 363–373 (1972).

[10] W. E. Blazier, ''Revised noise criterion for application in the acoustical design and rating of HVAC systems,'' Noise Control Eng. J. **162**, 64–73 (1981).

[11] S. S. Stevens, ''Calculation of the loudness of complex noise,'' J. Acoust. Soc. Am. **28**, 807–832 (1956).

[12] E. Zwicker, G. Flottorp, and S. S. Stevens, ''Critical band width in loudness summation,'' J. Acoust. Soc. Am. **29**, 548–557 (1957).

[13] L. L. Beranek, ''Balanced noise criterion curves, NCB,'' J. Acoust. Soc. Am. **86**, 650–664 (1989).

[14] H. F. Kingsbury, ''Review and revision of room noise criteria,'' Noise Control Eng. J. **43**, 65–72 (1995).

[15] N. W. M. Ko, W. F. Ho, and W. K. Un, ''Responses to air conditioning system noise,'' J. Sound Vib. **57**, 595–602 (1978).

[16] L. L. Beranek, W. E. Blazier, and J. J. Figwer, ''Preferred noise criterion (PNC) curves and their application to rooms,'' J. Acoust. Soc. Am. **50**, 1223–1228 (1971).

[17] ANSI S12.2, *American National Standard Criteria for Evaluating Room Noise* (The Acoustical Society of America, New York, 1995).

[18] N. Broner and H. G. Leventhall, ''A criterion for predicting the annoyance due to lower level low frequency noise,'' J. Low Freq. Noise Vib. **24**, 160–168 (1983).

[19] H. Tachibana, H. Yano, and Y. Sonoda, ''Subjective assessment of indoor noises—Basic experiments with artificial sounds,'' Appl. Acoust. **31**, 173–184 (1990).

[20] von E. Paulus and E. Zwicker, ''Computer programmes for calculating loudness from third-octave band levels or from critical band levels,'' Acustica **27**, 253–266 (1972).

[21] ISO 532, ''Acoustics—Method for Calculating Loudness Level'' (International Organization for Standardization, Geneva, 1975).

[22] R. M. Bethea and R. R. Rhinehart, *Applied Engineering Statistics* (Marcel Dekker, New York, 1991).

[23] W. R. Thompson, ''On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation,'' Ann. Math. Stat. **6**, 214–219 (1935).

[24] BS 8233, *Sound Insulation and Noise Reduction for Buildings* (British Standards Institution, London, 1987).

[25] D. D. Reynold, *Engineering Principles of Acoustics* (Allyn and Bacon, Boston, 1981).

[26] D. W. Robinson, ''Towards an unified system of noise assessment,'' J. Sound Vib. **14**, 279–298 (1971).

[27] L. L. Beranek, ''Application of NCB noise criterion curves,'' Noise Control Eng. J. **33**, 45–56 (1989).

Experimental study of different approaches for active control of sound transmission through double walls

C. Bao and J. Pan

Department of Mechanical and Materials Engineering, University of Western Australia, Nedlands, Western Australia 6907, Australia

Recently, active noise control technology has been used to increase the sound transmission loss of double-wall structures. Several approaches have been put forward and explored individually. However, no comparative study on those approaches has been conducted to show which approach is more effective for given circumstances. In this paper, three different approaches for controlling sound transmission through double-panel partitions into a room, i.e., applying acoustic control sources in the air gap between the two panels (cavity control), applying vibration control sources on the radiating panel (panel control), and applying acoustic control sources in the receiving room (room control), are studied and compared to each other experimentally. The mechanisms involved in each approach are illustrated and the conditions for effective noise attenuation are examined. The results show that the modal overlap of each subsystem (the cavity, the radiating panel, and the room) and the control mechanisms involved are the two most important factors that determine the effectiveness of each corresponding approach. © *1997 Acoustical Society of America.* [S0001-4966(97)05408-8]

PACS numbers: 43.50.Ki, 43.55.Rg [GAD]

INTRODUCTION

Double-panel partitions are widely applied in noise control engineering when relatively high values of the acoustic transmission loss have to be realized with lightweight structures. Examples include mobile office partitions, partition walls in buildings, and aircraft fuselage shells. In the low-frequency band, such partitions are, however, less efficient due to the resonance effect of double-panel systems (panel–cavity–panel coupling systems). Since active noise control is mostly efficient at lower frequencies, it is worthwhile to explore its potential to increase the transmission loss of double-panel partitions. Recently a number of investigations have been conducted in this regard. These investigations can be grouped into three approaches in terms of types and locations of control sources.

The first approach is referred to as cavity control in this paper. It attenuates sound radiation into a room by means of controlling the sound field in the air gap (the cavity) between the panels by using acoustic sources in the cavity. This approach is represented by the work of Sas and Bao (1993) and Sas *et al.* (1995). They have studied cavity control theoretically, numerically, and experimentally and have shown that it is very effective for certain cases. The increase of transmission loss in their experiment was as high as 40 dB. Similar work has also been conducted by Grosveld and Shepherd (1991), and by Gagliardini and Bouvet (1993). In the work of Grosveld and Shepherd, experiments were carried out on a mock-up of an aircraft fuselage. Global attenuation was achieved in an area of 50×50 in.^{2} for a single-frequency noise by using four control loudspeakers. In the work of Gagliardini and Bouvet, the control of broadband noise was considered. They were able to achieve an increase of transmission loss of 10 dB over the frequency range between 80 and 160 Hz.

The second approach is active structural acoustic control (ASAC) at the partition panels (referred to as panel control). In this case, the radiated sound is attenuated by means of controlling the vibration of the radiating panel using force sources mounted on the panel. This approach is represented by the work of Caneal and Fuller (1993). In their work, control moments, generated by piezoelectric actuators, were applied to the panels to minimize the radiated acoustic field into a room. An increase of 20 dB in transmission loss was achieved for certain cases. Similar work has also been conducted by Thomas *et al.* (1992). They achieved a reduction of sound pressure by 40 dB.

The third approach is referred to as room control that attenuates the radiated sound by means of acoustic sources in the receiving room. Unlike the other two approaches, this approach is often not viewed as a method of controlling sound transmission but rather as an application of active noise control in enclosures or in free space.

So far, those approaches have been explored individually but no comprehensive comparison has been published to show which approach is more effective for given circumstances. Thus the main objective of our study is to have a systematic comparison among different approaches. The study is conducted analytically and experimentally. The result of the analytical study has been presented elsewhere (Pan and Bao, 1996), showing the mechanisms involved in each approach and conditions for their effective application. The result of the experimental study is presented in this paper, where comparisons of the different approaches are made in terms of different criteria such as the frequency band of noise reduction, the change of panel vibration and the noise reduction in the area close to the partition, and the effect of thermal insulation in the cavity.
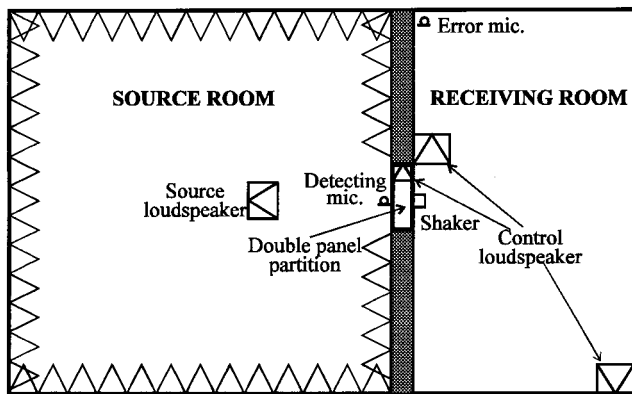
FIG. 1. Schematic presentation of the experimental setup.

I. EXPERIMENTAL SETUP

In the experiment, a double-panel partition is mounted in a common wall between an anechoic chamber (the source room) and a control chamber (the receiving room). The arrangement of the test chambers and the double-panel partition is shown in Fig. 1. The volume of the receiving room is 56 m^3 and its reverberation time below 200 Hz is around 1.7 s.

The double-panel partition consists of two aluminum panels of 2-mm thickness, separated by an air cavity of 275-mm depth. The other two dimensions of the air cavity are 2150 and 900 mm, respectively. The side walls of the cavity are concrete. The two panels are clamped to two heavy steel frames, respectively, allowing no vibration transmission between the two panels. The unclamped area of the panels is 2035×780 mm^2. The incident panel is excited by the loudspeakers placed in the anechoic chamber and 1.5 m away from the panel. Once excited, the incident panel radiates energy into the air cavity, thereby exciting the radiating panel, which in turn radiates energy to the receiving room. Unless otherwise stated, harmonic excitation is used in all the experiments described, as it serves better than broadband noise in revealing the control mechanisms involved in each approach.

An adaptive feedforward controller is used for the control. The reference signal of the controller is taken from a detecting microphone placed in the source room close to the partition, and the error signals from error microphones placed either in the receiving room or within the partition. As indicated in the analytical study, the quantitative understanding of the control mechanisms and performance comparison can be obtained from systems with any number of control sources. Thus for simplicity, a single control source is used in all the experiments except for the case where the effect of thermal insulation is investigated. For cavity control, the control loudspeaker is placed in a corner of the cavity. In panel control, a minishaker is mounted at the center of the radiating panel as the control source. As for room control, two locations are selected for the control loudspeaker: one is next to the radiating panel and the other in the far corner of the receiving room. A single error microphone is usually placed in a corner of the receiving room. Occasionally, two error microphones in other locations are also used.

II. COMPARISON IN TERMS OF FREQUENCY BAND OF REDUCTION

In this set of experiments, the performance of each control arrangement is evaluated by measuring the controlled and uncontrolled sound pressure. The sound-pressure level (SPL) was measured at 15 selected points distributed over the receiving room and processed into two measures to provide the basis for the evaluation. One measure is the number of locations where SPL increases. This indicates if the noise reduction is global. It is regarded as the global reduction if this number is less than three. The other measure is the averaged reduction level (ARL) derived from comparison of the averaged SPL over 15 locations with and without control [i.e., ARL$=10 \log_{10}(\Sigma p_u^2/N) - 10 \log_{10}(\Sigma p_c^2/N)$, where p_c and p_u are sound pressure with and without control, respectively, and N is the number of measurement locations]. This measure indicates the amount of reduction.

A. Room control

The experiments described in this section are designed to reveal the mechanism involved in room control. As suggested in the analytical study (Pan and Bao, 1996), the necessary condition for the global reduction of room control is that the two modal vectors due, respectively, to the primary (the radiating panel) and secondary (the loudspeaker) sources are mathematically proportional. The above condition can only be satisfied in two situations. One is to place the secondary source as close as possible to the primary source, allowing a similar generation of all dominated modes. The other happens at those frequencies where the modal overlap is low: i.e., the sound field is dominated by a single room mode and the contribution of nonresonant modes is small. In the latter situation, it is not necessary to have the secondary source close to the primary source as long as the secondary source can excite the same modes as the primary source does.

In the experiments, both situations are observed. In one of the control arrangements, the loudspeaker is placed next to the radiating panel (the next-to-source control). In the other arrangement, the control loudspeaker is in the far corner of the room (the away-from-source control), which enables excitation of all the room modes. Table I summarizes the experimental results. As expected, the away-from-source control is indeed quite effective at 26 Hz where the sound field is dominated by the first room mode. A noise reduction of 15 dB is achieved. Due to the damping of the room and the close spacing between the higher room modes (which results in a higher modal overlap), the sound field after 29 Hz is never clearly dominated by a single mode. Consequently, the away-from-source control is only effective in the low-frequency range up to 29 Hz. Whereas for the next-to-source control, since the control loudspeaker is very close to the radiating panel (which allows a similar generation of the sound field up to certain frequencies) the controllable frequencies extend up to 64 Hz. Due primarily to the different nature of the primary (panel) and secondary (loudspeaker) sources and also to the inevitable distance between them, the similarity between the two corresponding vectors is quickly diminished as the frequency increases. Consequently, global

TABLE I. Comparison between the two room control arrangements. ARL: averaged reduction level over 15 locations; NOI: number of locations where SPL increases.

Frequency (Hz)		26	29	51	59	62	
Next-to-source control	ARL(dB)	**15.5**	**5.0**	**14.1**	**0.9**	**10.6**	No global control after 64 Hz
	NOI	2	1	1	11	0	
Away-from-source control	ARL(dB)	**15.1**	**4.4**		No global control after 29 Hz		
	NOI	1	1				

control is no longer achievable at higher frequencies, even with the next-to-source control. It is interesting to note the control spill-over at 59 Hz where the modal overlap is not higher than that at 62 Hz. The modal analysis of the room indicated that the room modes that were excited by the panel at 59 Hz could not be properly excited by the control loudspeaker due to its location. This again reveals that the mechanism of room control is purely modal suppression.

In summary, the experiments described in this section confirm that the mechanism of room control *is* modal suppression. In order to have the maximum control effect, the control loudspeaker should be placed as close as possible to the radiating panel.

B. Cavity control

One obvious mechanism of cavity control is to break the noise transmission path: i.e., to attenuate the sound field in the cavity. This reduces the vibration of the radiating panel and consequently the sound radiation into the receiving room. This mechanism relies on modal suppression of the cavity modes. Another mechanism can also be involved in cavity control: i.e., to change the vibration pattern of the radiating panel through modal rearrangement of the cavity modes to form a weaker sound radiator. This is possible as the vibration pattern of the radiating panel is largely determined by the sound field in the cavity. Our analytical study (Pan and Bao, 1996) has indeed shown that each pressure component in the receiving room can be expressed as the linear combination of the pressure modal components of the cavity. This suggests that the sound pressure in the room can be attenuated by the rearrangement of the cavity modes. Therefore one of the aims of this section is to experimentally confirm the two mechanisms which apply to cavity control. In addition to the SPL at the 15 locations in the receiving room, the SPL at six selected locations in the cavity and the vibration level at six selected positions on the radiating panel are also measured. In the experiments, four different error microphone arrangements are used. One is referred to as the external sensing where the microphone is placed outside of the cavity, in a corner of the receiving room. The other three are referred to as the internal sensing where the microphones are placed inside the cavity. They include use of a single microphone at two different locations (internal sensing I and II) and use of two error microphones simultaneously (internal sensing III). While the internal sensing system is more compact and less sensitive to any change in the receiving room, the external sensing system may have better performance as the error signals are related more directly to the ultimate aim of the global noise reduction in the room.

Some typical results from the experiments are summarized in Table II, where the averaged noise reduction in the room as well as in the cavity and the change in the panel vibration are listed for each arrangement. From the experiments, cavity control is effective in the frequency range up to 102 Hz. This range can be divided into three subranges to show the control mechanisms involved.

In the lower-frequency range up to 58 Hz, the global reductions are achieved both in the cavity pressure and in the panel vibration (see the result at 51 Hz in Table II). Thus the control mechanism involved for the noise reduction in the receiving room is clearly modal suppression. The modal analysis of the cavity shows that the sound field in this frequency range is controlled by the (0,0,0) mode. Consequently, the reduction achieved in the room is not sensitive to the location of the error microphone, and very good reduction in the receiving room is achieved by using both internal (26 dB at 51 Hz) and external (26 dB at 51 Hz) error sensing.

From 58 to 80 Hz, the global reduction is achieved in the cavity pressure but not in the panel vibration (see the result at 62 Hz in Table II). This is quite surprising, as it implies that the global reduction of the sound pressure in the cavity does not necessarily lead to the global reduction of the panel vibration. The modal analysis of the cavity shows that the sound field is now dominated by the first cavity mode. Thus global noise control in the cavity is expected by using a single control source with modal suppression. As other cavity modes also have some influence over this frequency range, the controlled (residual) field is still strong and complex enough to excite the previously unexcited vibration modes. This explains why the global reduction in the cavity

TABLE II. Comparison among different arrangements of cavity control. ARL: averaged reduction level; PVR: panel vibration reduction.

Frequency (Hz)		51	62	100
External sensing	ARL (room, dB)	**26.2**	**13.2**	**9.3**
	ARL (cavity, dB)	11.6	6.0	2.7
	PVR	yes	no	no
Internal sensing I (single microphone)	ARL (room dB)	**26.1**	**14.0**	**−2.5**
	ARL (cavity, dB)	13.8	4.9	−0.8
	PVR	yes	no	no
Internal sensing II (single microphone)	ARL (room, dB)	**25.2**	**6.8**	**10.6**
	ARL (cavity, dB)	8.2	7.1	−1.8
	PVR	yes	no	no
Internal sensing III (two microphone)	ARL (room, dB)	**25.8**	**8.9**	**5.6**
	ARL (cavity, dB)	8.3	7.7	1.4
	PVR	yes	no	no

pressure does not lead to the global reduction in the panel vibration. Also because of the participation of other cavity modes, different distributions in the cavity pressure now result from different error microphone arrangements. Moreover, these different pressure distributions correspond to different reduction levels in the room (e.g., 14-dB reduction of internal sensing I compared to 7-dB reduction of internal sensing II, at 62 Hz). The result clearly suggests that modal rearrangement is involved. The control mechanism in this case is really the combination of modal suppression and modal rearrangement, with modal rearrangement playing a more important part. It should be noted that the external sensing system can always achieve a high level of reduction (e.g., 13 dB at 62 Hz).

From 80 to 102 Hz, the global reduction cannot be achieved in both the cavity pressure and the panel vibration (see the result at 100 Hz in Table II). For some frequencies, reduction in the room is not even achievable. The modal analysis shows that the cavity field is now dominated by at least two cavity modes. The reduction in the room is clearly due to the rearrangement of those cavity modes, which changes the radiating panel into a weak radiator. As for the internal sensing, unlike in the lower-frequency range, the panel vibration patterns with different error microphone locations can be very different. Consequently, the reduction in the room is very sensitive to the location of the error microphone. The arrangement that produces a similar pressure distribution in the cavity as that of the external sensing yields the largest reduction, while those with different distributions from that of the external sensing achieve no reduction or even increase the noise level in the room.

In summary, the two mechanisms involved in cavity control are clearly shown in the experiments. In terms of the error microphone arrangement, the external error sensing gives good results over the whole controllable frequency range as it can always choose the most appropriate mechanism. As the internal error sensing aims at attenuation of the cavity pressure, it is naturally associated only with the mechanism of modal suppression. Therefore it works well at low frequencies where global attenuation of the cavity pressure is obtainable. However, good results cannot be obtained by this arrangement at higher frequencies where modal rearrangement is actually required.

C. Panel control

It is well known that two mechanisms are involved in panel control (Pan, 1988; Fuller *et al.*, 1991). One is modal suppression which suppresses the panel vibration thereby reducing the radiated sound. For a single-channel control, modal suppression is often associated with the attenuation of sound transmission near the resonance frequency of a panel-controlled mode where the transmitted energy is carried mainly by a single-panel mode. The other mechanism is modal rearrangement which rearranges the magnitudes and phases of the panel modes thereby changing the radiating panel to a weaker radiator. This can often lead to an increase of overall panel vibration. Modal rearrangement is associated with the sound transmission into the room-controlled modes where the energy is carried by several panel modes.

TABLE III. Typical results of panel control. ARL: averaged reduction level over 15 locations; NOI: number of locations where SPL increases.

Frequency (Hz)	26	29	51	62	74	80	85	110	116	132	155
ARL (dB)	14.5	8.2	27.9	12.7	9.9	3.2	4.7	1.8	4.2	1.1	6.1
NOI	0	0	0	0	0	4	1	5	2	7	3

Some typical experimental results from panel control are listed in Table III. It can be seen that the controllable frequency range of panel control can go up to 155 Hz. This is because the modal density of the panel does not increase as the frequency increases. Therefore the modal overlap of the panel (the major factor that determines how high the frequency of the modal control can be) depends only on the damping factor which increases with the frequency. The controllable frequency range can further be divided into two subranges. Up to 42 Hz, the global reduction in the room is always achievable, thanks to a low modal overlap of the panel. Beyond that, the reduction can be achieved only at certain discrete frequencies where the modal overlap is low (usually at the resonance frequencies of the panel). In the higher-frequency range, the performance is sensitive to the location of the error microphones. The experiment in this regard shows that using multiple error microphones can avoid the problem.

D. Comparison of the three approaches

As demonstrated in the previous sections, the next-to-source arrangement in room control and the external sensing arrangement in cavity control yield better results than other arrangements of these two control approaches. Thus hereafter when room control and cavity control are mentioned, the results from those two arrangements are used for the comparison.

Table IV summarizes the typical results of the three control approaches. It can be seen that cavity control and panel

TABLE IV. Typical results of the three approaches. ARL: averaged reduction level over 15 locations; NOI: number of locations where SPL increases; BB35−55: bandlimited noise from 35 to 55 Hz; BB45−85: bandlimited noise from 45 to 85 Hz.

Frequency (Hz)	Room control ARL (dB)	NOI	Cavity control ARL (dB)	NOI	Panel control ARL (dB)	NOI
BB35−55	7.6	0	9.7	0	1.6	4
BB45−85	2.5	4	6.3	0	0.2	7
26	15.5	1			14.5	0
48	8.0	4	17.9	1	15.1	0
51	14.1	1	26.2	0	27.9	0
55	13.3	1	14.8	0	15.5	0
62	10.6	0	13.2	0	12.7	0
74			16.2	0	9.9	0
80			6.0	0	3.2	4
85			2.4	1	4.7	1
89	No reduction after		1.9	5	2.6	3
100	64 Hz		9.3	1	2.2	2
116					4.2	2
155			No reduction after		6.1	3
			102 Hz		No reduction after	
					155 Hz	

TABLE V. Typical near-field noise reductions. ARL: averaged reduction level in dB over 9 locations; NOI: number of locations where SPL increases.

Frequency (Hz)	Cavity control						Panel control					
	3 cm		10 cm		36 cm		3 cm		10 cm		36 cm	
	ARL	NOI	ARL	NOI	ARL	NOI	ARL	NOI	ARL	NOI	ARL	NOI
26							4.3	3	9.5	2	14.2	0
51	19.0	0	19.0	0	19.5	0	1.2	3	8.9	0	20.5	0
62	11.2	0	12.9	0	15.6	0	7.8	0	10.2	0	12.3	0
74	2.3	2	6.1	0	11.4	1	2.3	3	5.8	1	7.5	0
85	0.3	5	3.8	4	6.0	2	−3.1	8	−2.6	8	−1.7	7
100	3.1	4	4.9	3	11.0	1						
116							2.3	3	4.2	1	3.2	1

control are much better than room control in terms of the global reduction, averaged reduction level, and the frequency band of reduction. This can be explained as follows. First, both cavity and panel control are applied to the transmission path, while room control is applied to the radiated sound in the room. Better global control and a larger amount of reduction can be expected when the transmission path is effectively controlled. Second, both cavity and panel control work on the basis of the two mechanisms, modal suppression and modal rearrangement, while room control is on the basis of the single mechanism, modal suppression. Finally, the modal overlap of the room is much higher than that of the cavity due to the much larger size of the room, and also higher than that of the panel at higher frequencies due to the fact that the modal density of the panel does not increase with frequency while that of the room does. The latter two reasons result in cavity and panel control achieving noise reduction over a wider frequency range.

Hereafter, the focus of the comparison will be between cavity control and panel control. For this case, the comparison is based on the frequency band of reduction, the averaged reduction level, and the effect on the panel vibration.

Although panel control is effective at higher frequencies (up to 155 Hz) due to constant modal density of the panel, the global noise reduction can only be achieved at discrete frequencies. On the other hand, cavity control yields the global reduction continuously up to 85 Hz. Thus cavity control is effective in attenuation of broadband noise and panel control is not. In this respect, panel control is even worse than room control, as shown by the first two rows in Table IV. This can be explained as follows. Panel control relies mainly on modal rearrangement as the panel vibration is normally composed of more than one panel mode at most frequencies. If a point control source is applied, its effectiveness is very sensitive to the source location. For the fixed source location, it is effective only at discrete frequencies. While for cavity control, the cavity pressure is dominated by either the (0,0,0) mode or the first cavity mode up to 80 Hz. Thus a single loudspeaker in an appropriate location is effective over that whole frequency range. Also as cavity control applies the pressure force on the whole area of the panel instead of a single point, its effectiveness is less sensitive to the excitation (noise) frequencies. In summary, for control of broadband noise or the tonal noise with a variable frequency, cavity control is superior to panel control.

The noise reduction level achieved by panel control is as high as that of cavity control at controllable frequencies, as both approaches are applied to the transmission path and modal rearrangement is as effective as modal suppression.

As panel control relies on modal rearrangement at most frequencies, the noise reduction in the receiving room is often at the expense of an increase of the panel vibration. For cavity control, the panel vibration is always reduced in the low-frequency range up to 58 Hz, because modal suppression is the mechanism in this frequency range. Beyond that frequency range, an increase of panel vibration level can often be observed. Thus in terms of the effect on the panel vibration, cavity control does better than panel control at lower frequencies and has a similar behavior at higher frequencies.

III. COMPARISON IN TERMS OF OTHER ASPECTS

As demonstrated in the previous section, both the cavity and panel control can achieve better noise attenuation than room control in many aspects. Therefore the comparison in this section is conducted only between the cavity and panel control approaches.

A. Near-field reduction

The effect of control on the sound field close to the panel is important for certain applications such as in aircraft where passengers have to sit close to fuselage trim panels. To investigate this effect, sound-pressure measurements were conducted at a distance of 3, 10, and 36 cm from the panel, respectively. For each distance, SPLs were measured at nine points in a 3×3 array on a 60×120-cm rectangular area.

Table V summarizes the typical results of the experiments. It should be noted that the global noise reductions in the room were achieved at all frequencies listed in the table. From the experiments, the noise reduction in the area close to the radiating panel (near-field reduction) is closely related to the control mechanism involved. For those frequencies where modal suppression is involved, the global reduction can be achieved even at the distance of 3 cm from the panel, as shown by the results obtained in the low-frequency range up to 58 Hz in cavity control. On the other hand, if the control mechanism involved is modal rearrangement, the near-field reduction cannot be guaranteed anymore. In this

FIG. 2. Measured relative sound index in the receiving room without (----) and with (———) the thermal insulation.

TABLE VI. Effect of thermal insulation on cavity control. ARL: averaged reduction level in dB over 15 locations; NOI: number of locations where SPL increases; BB40–80; bandlimited noise from 40 to 80 Hz.

Frequency (Hz)	Single-channel control				Two-channel control			
	Insulation		No insulation		Insulation		No insulation	
	ARL	NOI	ARL	NOI	ARL	NOI	ARL	NOI
BB40–80	**7.7**	0	**7.5**	0				
100	**9.3**	1	**9.3**	1				
115	**−2.7**	9	**−2.5**	9	**4.9**	1	**5.2**	0
159					**1.2**	6	**8.6**	2
166					**0.8**	7	**7.1**	1

control yields better results in terms of the global noise reduction in the near field.

B. Reduction with thermal insulation

In many double-wall applications, thermal insulation materials are placed in the gap between the two walls. The effect of thermal insulation on both cavity and panel control has been investigated experimentally. In the experiments, a 100-mm-thick polyester blanket was placed in the gap of the double panels. The density of the material is 7.2 kg/m^3 and it is often used in the building industry.

Figure 2 shows the relative transmitted sound energy index (the sound-pressure square averaged over 15 points) in the receiving room with and without the thermal insulation. It can be seen that the thermal insulation has little effect on the sound transmission up to 220 Hz in general. At a few resonance frequencies it has some marginal influence: about 4-dB reduction in the frequency range of 45–60 Hz, 5 dB around 80–100 Hz, 8 dB around 165–185 Hz, and 4 dB around 190–210 Hz.

Figure 3 shows the relative sound energy index in the cavity (excited by the control loudspeaker and averaged over six points) with and without the thermal insulation. In this case, the thermal insulation has more influence, especially after 160 Hz. The resonances of the cavity are clearly damped out, which implies an increase of the modal overlap. Thus it is expected that cavity control will be less effective at higher frequencies. Since the single-channel control can only work in the frequency range up to 102 Hz, the two-channel control has to be used in order to verify that prediction.

case, the near-field reduction is very much dependent on how the panel vibration is rearranged. For some rearrangements (e.g., at 62 Hz in panel control), the near-field reduction is still achievable. While for some other rearrangements (e.g., at 85 Hz in panel control), the near-field reduction cannot be achieved. In the experiments, cavity control tends to give a better result when the mechanism of modal rearrangement is involved. For example, at 85 Hz the global reduction can be achieved to some extent at the distance of 36 cm from the panel by cavity control but not by panel control (in fact, SPL has been globally increased). This indicates that the rearrangement of the panel modes through cavity control may be more effective, as it applies the pressure force on the whole panel instead of at some discrete points. For this reason and also due to its wide frequency range of modal suppression, cavity control has achieved the global reduction at the distance of 3 cm in the frequency range up to 58 Hz, at 10 cm up to 80 Hz, and at 36 cm in the whole controllable frequency range (up to 102 Hz). While for panel control, even at as low as 26 Hz the global reduction is not possible at the distance of 3 cm. At the distance of 10 cm, the global reduction can be achieved in the frequency range up to 74 Hz. As the frequency increases, even at the distance of 36 cm the global reduction is not possible. The experiments also show that the near-field reduction is very sensitive to the location of the control source for panel control. In summary, cavity

FIG. 3. Measured relative sound index in the cavity without (----) and with (———) the thermal insulation.

FIG. 4. Measured relative vibration index on the radiating panel without (----) and with (———) the thermal insulation.

TABLE VII. Effect of thermal insulation on panel control. ARL: averaged reduction level over 15 locations; NOI: number of locations where SPL increases.

Frequency (Hz)		26	51	74	116	155
Insulation	ARL(dB)	**14.1**	**22.5**	**9.9**	**3.8**	**5.9**
	NOI	0	0	0	2	2
No insulation	ARL(dB)	**14.3**	**25.6**	**9.8**	**4.1**	**6.2**
	NOI	0	0	0	2	3

Table VI summarizes the experimental results of the various control arrangements. It does show that the thermal insulation has little effect on cavity control below 130 Hz. After 130 Hz, it has a negative effect on cavity control. Without thermal insulation the two-channel control is effective up to the frequency of 200 Hz while with the insulation it is only effective up to 138 Hz.

Figure 4 shows the relative vibration index on the radiating panel (exited by the loudspeakers in the source room and averaged over six points) with and without the thermal insulation. It can be seen that the thermal insulation has some effect on the vibration index as the resonant peaks are attenuated by 2–5 dB. This is due to the similar attenuation of sound pressure in the cavity. However, as the modal characteristics (e.g., modal overlap) of the panel vibration change a little, it is expected that the thermal insulation has little effect on panel control. This is confirmed by the experiment and the results are shown in Table VII.

In summary, the thermal insulation normally has little effect on the sound transmission at low frequencies. While the thermal insulation has little effect on panel control due to its minimal effect on the modal characteristics of the panel vibration, it does have a negative effect on cavity control due to an increase of the modal overlap of the cavity.

IV. CONCLUSIONS

Three different approaches (i.e., cavity control, panel control, and room control) in active control of sound trans-mission through double-panel partitions have been investigated and compared to each other experimentally.

It has been shown that room control is not as effective as cavity and panel control methods if the double panel is the only sound transmission path. In such cases, room control should be regarded as the last option when applying active control for attenuation of sound transmission. As for the other two approaches, cavity control normally performs better than panel control in the areas of control of broadband noise, attenuation of the panel vibration, and control of the near-field sound radiation.

The experiment has also shown that the thermal insulation has little effect on panel control. As for cavity control, although the thermal insulation also has little effect at lower frequencies, it does reduce the control effectiveness at higher frequencies.

Caneal, J. P., and Fuller, C. R. (**1993**). ''Active structural acoustic control of noise transmission through double panel systems,'' 15th AIAA Aeroacoustic Conference, AIAA-93-4421.

Fuller, C. R., Snyder, S. D., and Hansen, C. H. (**1991**). ''Active control of sound radiating from a vibrating rectangular panel by sound sources and vibration inputs: an experimental comparison,'' J. Sound Vib. **145,** 195–215.

Gagliardini, L., and Bouvet, P. (**1993**). ''DAP: the active controlled double wall,'' Proc. Inter-Noise 93, Vol. I, 107–110.

Grosveld, F. W., and Shepherd, K. P. (**1991**). ''Active sound attenuation across a double wall structure,'' 29th AIAA Aerospace Science Meeting AIAA-91-0498.

Pan, J. (**1988**). ''A study of the active control of sound transmission through a panel into a cavity,'' J. Acoust. Soc. Am. Suppl. 1 **84,** S182.

Pan, J., and Bao, C. (**1996**). ''Active attenuation of noise transmission through elastic partitions with high modal densities,'' Proc. Inter-Noise 96, Vol. III, 1055–1060.

Sas, P., and Bao, C. (**1993**). ''Use of active noise and vibration control for improving insertion loss of double wall structures,'' J. Dutch Acoust. Soc. (NAG) **120,** 57–65.

Sas, P., Bao, C., Augusztinovisz, F., and Desmet, D. (**1995**). ''Active control of sound transmission through a double-panel partition,'' J. Sound Vib. **180,** 609–625.

Thomas, D. R., Nelson, P. A., Pinnington, R. J., and Elliott, S. J. (**1992**). ''Active control of sound transmission through stiff lightweight composite fuselage constructions,'' Proc. 14th AIAA Aeroacoustics Conference **II,** 552–560.

Active suppression of sound diffracted by a barrier: An outdoor experiment

A. Omoto, K. Takashima, and K. Fujiwara
Department of Acoustic Design, Kyushu Institute of Design, Fukuoka, 815 Japan

M. Aoki
Aoki Applied Acoustics, Kobe, 651-21 Japan

Y. Shimizu
Nichias Corporation, Tokyo, 105 Japan

The active control method was used to suppress the sound diffracted by an outdoor barrier. This method operated by the cancellation of the sound pressure at the diffraction edge of the barrier, which normally behaves like the virtual source of the diffracted field. The results of two experiments are shown in this paper. In the first experiment, we employed two independent controllers that utilized multi-channel adaptive signal processing to minimize the sum of the mean square of the sound pressure at four and six points along the diffraction edge. Measurement of sound pressure levels at various distances from the barrier showed effective sound suppression, with about 6-dB excess attenuation over the barrier's insertion loss at the receiver at a distance of 50 m. A practically realistic noise source, a fan blower, was used as a primary source in the second experiment and the moderate attenuation could be achieved at almost all the receiver points. The results obtained in these two experiments provided the experimental verification of the strategy for the active suppression of sound diffracted by a noise barrier. © *1997 Acoustical Society of America.* [S0001-4966(97)02009-2]

INTRODUCTION

An application of the active control technique for reducing the sound diffracted by an outdoor noise barrier is reported. The active control method is based on the cancellation of sound pressure at a diffraction edge, which behaves like a virtual source for the diffracted field.[1,2] The concept of a ''virtual source'' is only an interpretation of the approximate expression[3] of the far field from the edge; physically, the sound pressure at the vicinity of the edge has a dominant effect on the diffracted field. Ideal control of sound propagation would require the continuous cancellation of the sound pressure over an arbitrary region near the diffraction edge. In practice, however, such control is difficult to implement. Instead, our strategy focused on cancellation at multiple points along the diffraction edge. We conducted two experiments in a rice paddy, building the barrier outdoors to verify the validity of our application and to deduce additional obstacles to practical use.

In the first experiment, we assumed some ideal conditions, i.e., the noise source was a loudspeaker and the primary signal was completely (electrically) known. Two controllers which executed the multi-channel adaptive signal processing were used in this experiment. One controller utilized secondary sound sources at four equally spaced points along the edge of the barrier, and the other utilized six secondary sound sources. They worked simultaneously and independently to minimize the sum of the mean square of the sound pressure.

Therefore there are two important aspects to this experiment. The first is the effect of the active control: the excess attenuation in addition to the barrier's own sound shielding efficiency. The other is the behavior of the adaptive systems working simultaneously and independently.

The second experiment was carried out to verify the effect of our strategy in practically realistic situations. A fan blower was used as a primary noise source. In such cases we must know the signal which has high coherence with the primary noise to drive the secondary sources. The method of detection of primary signal and the effect of the control were examined.

I. EXPERIMENT WITH ELECTRICALLY KNOWN PRIMARY SIGNAL

A. Materials and method

The barrier was constructed of two 0.8-mm steel panels, which were separated by a 150-mm airspace packed with 150- kg/m^3 rock wool (Fig. 1). Erected in a rice paddy, the overall dimensions of the barrier were 2 m(H) ×20 m(W)×0.2 m(D). The transmission loss due to this barrier was measured at about 38 dB at 125 Hz.

Figure 2 shows the experimental setup, with the primary (noise) source located on the ground ($y=0$) at $z=0$, 4.0 m from the diffraction edge and at an angle of 60° from it. (The coordinate system was arbitrarily established as shown in Fig. 2.) The origin was set 1.0-m distant from the center of the barrier at ground level.

The two adaptive signal processing systems were designated ''System-1'' (4 channels) and ''System-2'' (6 channels). The secondary sources were located 0.5 m and at an
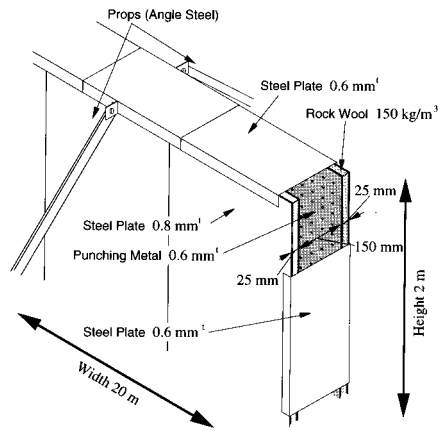
FIG. 1. Construction of the barrier.
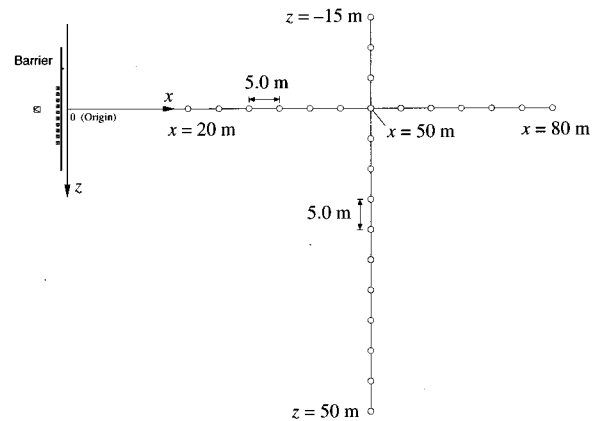
FIG. 3. The location of the distant receivers.

angle of 60° from the diffraction edge as shown in Fig. 2. They were spaced along the barrier at intervals of 1 m. Each point of cancellation was on the diffraction edge of the source side, and had the same z coordinate as the corresponding secondary source. An electric condenser microphone, whose diameter was 1 cm, was used as a error sensor located at each point of cancellation.

The near receiver area was comprised of a grid measuring 11×4 with 2-m spacing (area A in Fig. 2). More distant measurements were taken at heights of 1.0 m and 2.0 m at each point along the line $z=0$, at 5-m increments from $x=20$ to 80 m, and along the line $x=50$ m, at 5-m increments from $z=-15$ to 50 m (Fig. 3). Each location at which sound pressure measurements were taken was in the shadow region of the barrier, i.e., the primary source could not be detected from that position. In addition to these receivers, the distribution of the sound pressure along the diffraction edge of the barrier was measured at 0.2-m intervals to investigate the behavior of the adaptive systems (Fig. 4).
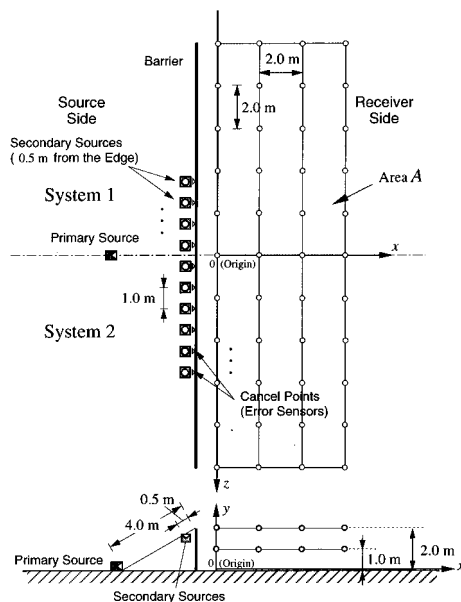
The primary source was driven with a 125-Hz one octave-band noise. In this experiment, this signal was also input to the signal processing controller to generate the signal which drives the secondary sources. In actual use, the primary sound would be detected by microphone or vibration pickup. This problem is discussed in the next experiment.

The sound pressure levels were measured at the receivers, first with the secondary sources turned off, and then again with the secondary sources producing the interference pattern. The effect of the active control was defined as the difference between the two sets of sound pressure measurements.

During the experiment, the weather was mild; there was only a light breeze. (The wind speed was 2–3 m/s.) There were therefore no serious atmospheric effects on the measurements or the behavior of the adaptive systems.

B. Adaptive signal processing

In our application, increasing the number of points of cancellation along the edge was the most effective way to expand the area of sound reduction. However, the number of single-controller points of cancellation is restricted by hardware limitations. Two independent controllers were therefore used to overcome these limits. The two systems worked simultaneously and independently to minimize the sum of the mean square of the sound pressure along the edge of the barrier. The convergence of the adaptive systems in such cases has been predicted.[2,4,5] We believe that using multiple controller systems may be the most practical approach when numerous channels are required.

FIG. 2. Experimental setup.
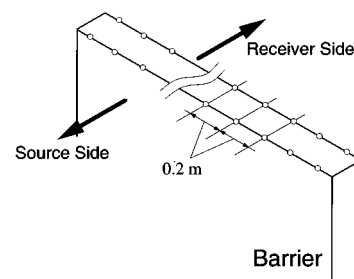
FIG. 4. Receiver locations on the diffraction edge of the barrier.
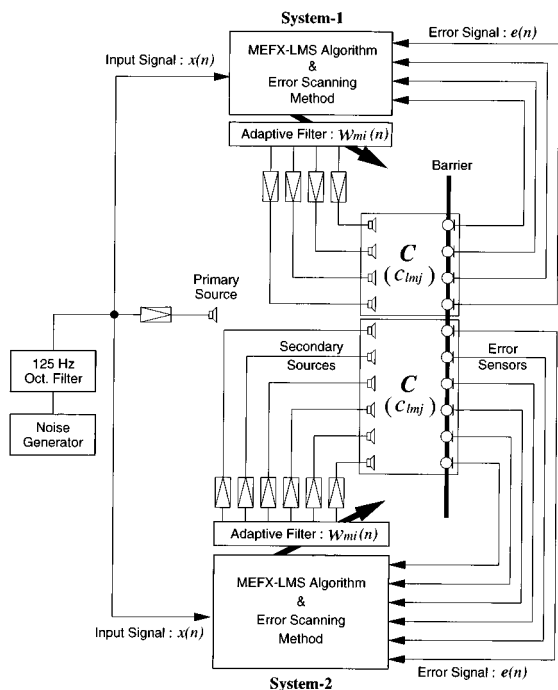
FIG. 5. Block diagram of two adaptive systems.

Figure 5 shows the block diagram of the system. Both controller systems used the MEFX-LMS[6] as their operating algorithm. Let C express the transfer matrix between the secondary sources and the error sensors for each system. The lth, mth element (impulse response between mth secondary source to lth error sensor) was modeled by the Jth order FIR digital filter. The adaptive filter, w, which modifies the amplitude and the phase of the secondary signal, is also modeled as Jth order FIR filter. The impulse responses from each secondary source to each point of cancellation (C_{lmj}) were measured by a conventional LMS algorithm prior to the ''real'' control process. The same procedure were done for systems 1 and 2 independently. The coefficients of the adaptive filter w were updated iteratively by the following equation,

$$w_{mi}(n+1) = w_{mi}(n) - \alpha \sum_{l=1}^{L} e_l(n) \sum_{j=0}^{J} C_{lmj} x(n-i-j),$$ (1)

where e_l is the output signal of the lth error sensor (located at the lth point of cancellation), x is the input signal to the system, α is the step size parameter, and n represents the sample number. L is the number of channels, i.e., $L=4$ for system 1 and $L=6$ for system 2, respectively. In an actual calculation of Eq. (1), the error scanning method[7] was introduced to reduce the amount of calculation. The sampling frequency was 1 kHz. The tap length of the adaptive filter and the FIR filter which generated the ''filtered-x'' signal were 50 taps in system 1, and 64 taps in system 2.

C. Results and discussion

Figures 6 and 7 show the distribution of the effectiveness of the control (dB) measured in area A in the case of (a) only system 1 (4-channels) functioning; (b) only system 2

(a) Only System-1 functioning

(b) Only System-2 functioning

(c) System-1 and 2 functioning
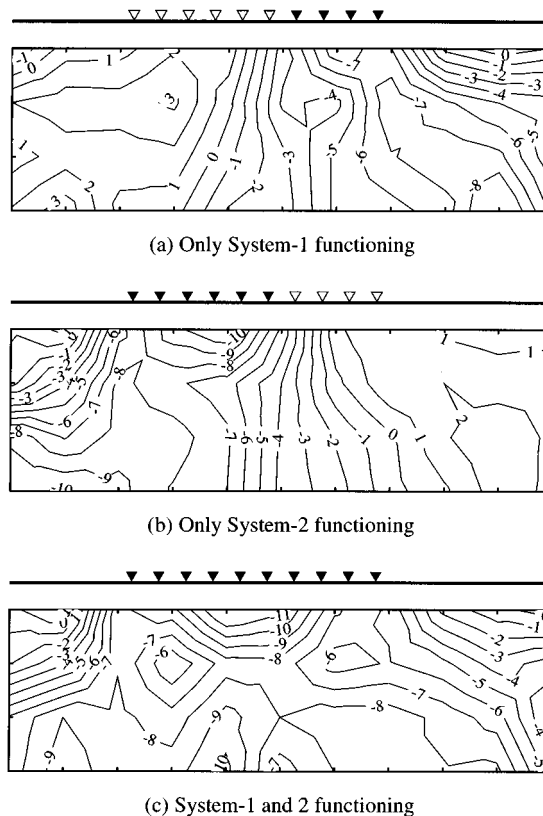
FIG. 6. Sound attenuation by active control in area A. ($y = 1$ m, ▼ indicates the units of the functioning control system.)

(a) Only System-1 functioning

(b) Only System-2 functioning

(c) System-1 and 2 functioning
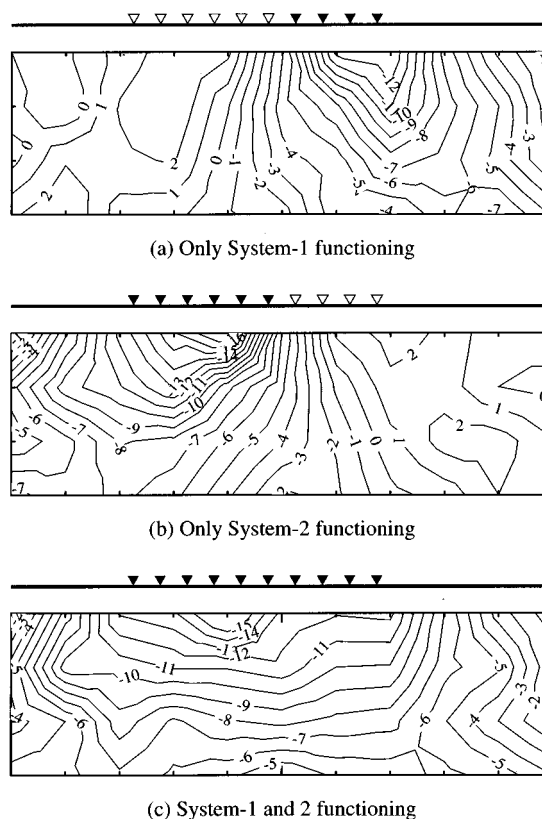
FIG. 7. Sound attenuation by active control in area A ($y = 2$ m).
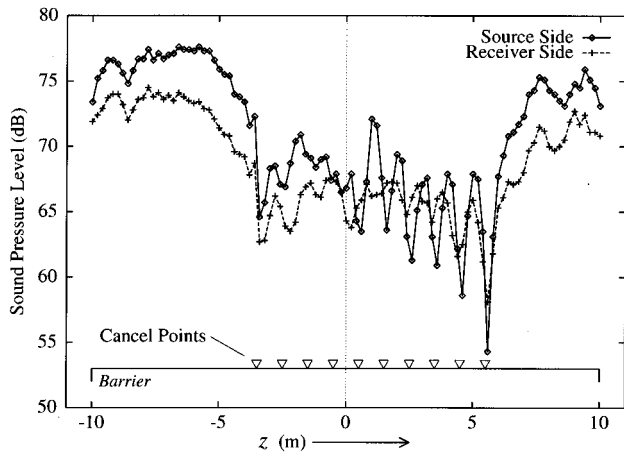
1673   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997   Omoto *et al.*: Active suppression of diffracted sound   1673

FIG. 8. The distribution of the sound pressure level along the edge of the barrier.

(a) $y = 1$ m

(b) $y = 2$ m

(c) Effect of the Active Control

FIG. 9. Sound pressure level at the distant receiver locations. (Along the line $z = 0$, $x = 20$ m to 80 m.)

(6-channels) functioning; and (c) with both systems functioning. The measurements for Fig. 6 were taken at a height of $y = 1$ m, and those for Fig. 7 were taken at a height of $y = 2$ m. A negative measurement in the figures indicates effective attenuation, and a positive measurement indicates amplification by the active control.

With only one system functioning, the greatest attenuation was attained near the barrier and toward the center of the functioning system. System 2 gave a greater maximum attenuation than system 1, as seen in Figs. 6 and 7. This divergence seems to be due to the difference in the number of channels and in the tap length of the adaptive filters. The attenuation gradually dropped off with increased distance from the functioning system, and 1–3 dB of amplification occurred at some of the more distant points shown in Fig. 6, and 1–2 dB of amplification occurred at those shown in Fig. 7.

There are 5–6 dB differences in the maximum attenuation between Figs. 6 and 7 ($y = 1$ m and $y = 2$ m). The sound pressure in the area of $y = 1$ m is fairly small as a result of the shielding efficiency of the barrier itself, and the control may further decrease this pressure. Even a weak reflected sound from the ground can therefore affect the sound pressure at such a sensitive (controlled) sound field. We postulate that the reflected sound from the ground is one of the reasons for decreased attenuation.

Figures 6(c) and 7(c) show that, with both systems functioning, more than 6 dB of attenuation was attained at most of the receiver points. The simultaneous operation of the two systems extended the area of effective attenuation. As a reference, Fig. 8 shows the distribution of the sound pressure level along the edge of the barrier. The two lines on the chart correspond to the sound pressure measurements taken at the edge of the barrier on both the source side and the receiver side. Six remarkable dips were noted; they represent the points of cancellation for system 1. The four dips of system 2 are not as clear. This difference in attenuation seems to correspond to the difference between the maximum attenuations in Figs. 6 and 7, (a) and (b). The sound pressure at the middle of the two systems was moderately suppressed.
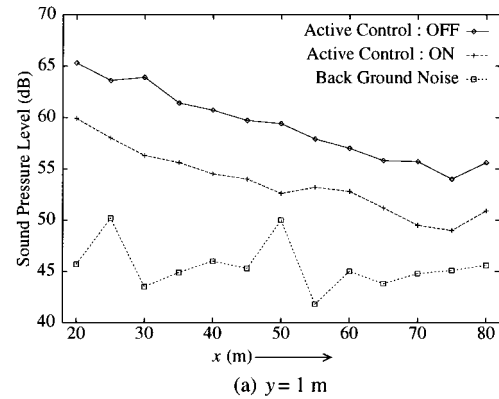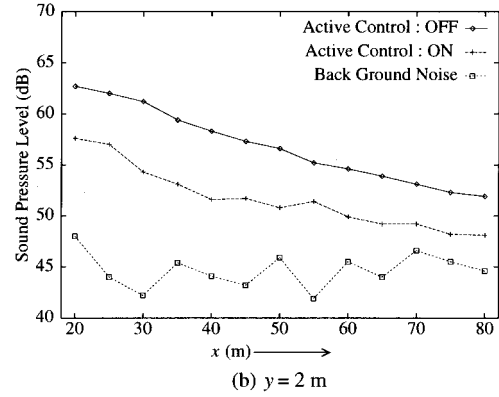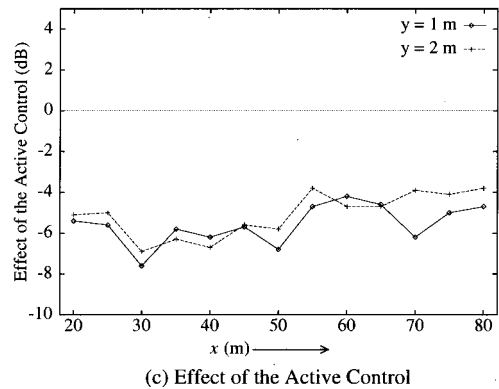
This observation shows that the algorithm can converge and that attenuation can be achieved in this case. This behavior of the algorithm was predicted theoretically in Refs. 2, 4 and 5. The results also suggest that the area of sound reduction could be extended by the use of more than two adaptive systems.

Figures 9 and 10 show the sound pressure level measured at the more distant receiver locations described in the previous section, corresponding to the receivers along the line $z = 0$ and $x = 50$ m. In each figure, (a) and (b) correspond to heights $y = 1$ m and 2 m, respectively, and (c) is the effect of the active control system. The three lines plotted in charts (a) and (b) show the sound pressure with the active controls turned off, that with both active controls turned on, and that of the background noise alone. In Fig. 9(a) and (b), the sound pressure level with the active system turned off can be seen to decrease gradually as x increases. The same

(a) $y = 1$ m

(b) $y = 2$ m

(c) Effect of the Active Control
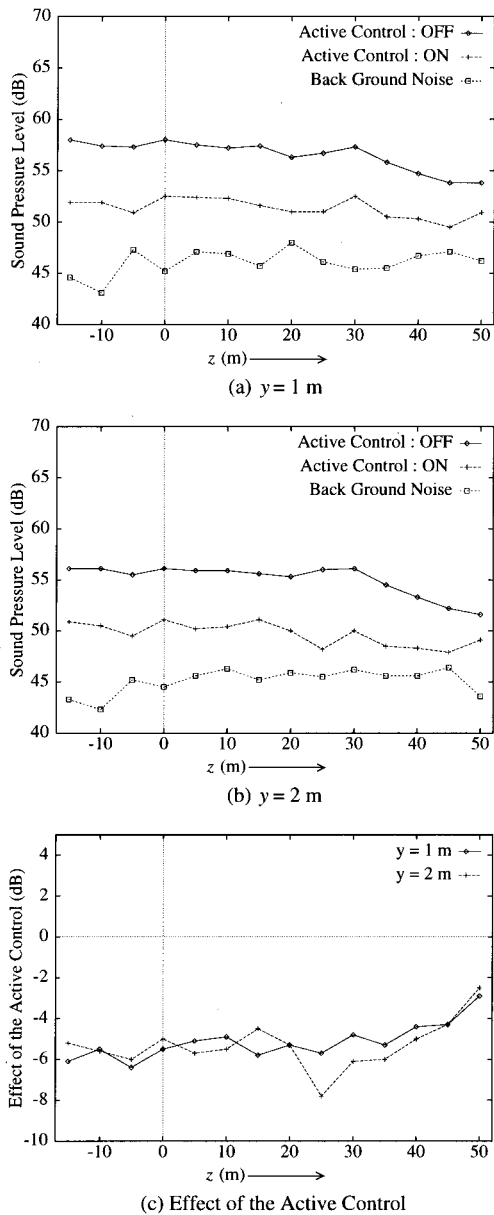
FIG. 10. Sound pressure level at the distant receiver locations. (Along the line $x = 50$ m, $z = -15$ m to 50 m.)

FIG. 11. A fan blower used in the experiment as a noise source.

II. EXPERIMENT WITH PRACTICAL NOISE SOURCE

A. Materials and method

The reference signal, which is used to produce the control signal, plays an important role in the performance of the active control system. In the previous experiment, this signal was assumed to be an identical electrical signal which was supplied to the primary source. In most actual cases, however, this signal is hardly available and some form of acoustic detector is generally needed to sense the noise coming from the primary source. We therefore conducted another experiment with a real noise source to simulate much more practical situations.

A fan blower (Fig. 11) was used as a primary noise source. The fan has eight blades and the blade passage noise was dominant. In the experiment, the fan was rotated at the speed of 15.6 r.p.s. and caused periodic noise at around 125 Hz. Figure 12 shows the power spectrum of the primary signal observed at the diffraction edge of the barrier used in the experiment.

Figure 13 shows the experimental setup. The barrier used in the previous experiment was again constructed but the length was 15.8 m in this case. There were some huts near the barrier. Additional barriers were then attached to both sides of the barrier to suppress the effect of the reflected sound from the huts. The effects of the control were measured at heights of 1.0 m and 2.0 m at each receiver along the line $z = -4$ and 4 m at 2-m increments and along the line $z = 0$ m at 1-m increments. Maximum x length at each z line was limited by the locations of the other huts.

trend occurs in the plot corresponding to that in which the active system is turned on, but diminished by 5–6 dB in the region $x < 50$ m, and by 4–5 dB in the region $x > 50$ m. The decreased attenuation in the region of $x > 50$ m seems to be due to the smaller signal to background noise ratio. In Fig. 10, an attenuation of 5–6 dB can be seen at all the receiver points except near $z = 50$ m, where it was less marked. Ignoring the background noise, the attenuation was more than 6 dB over most of the measurement area.

The stable attenuation at the distant receivers can be understood as follows: The diffracted field at distant receivers was far enough away from the barrier that the sound impacting on the receivers was effectively expressed only by the virtual source at the diffraction edge. The cancellation of the sound pressure at this virtual source had the same effect as reducing the radiation power at the sound source.

FIG. 12. Power spectrum of the primary signal observed at the diffraction edge.

The control system and the adaptive algorithm were identical to system 2 (6-channels) used in the previous experiment. The secondary sources were again located 0.5 m and at an angle of 60° from the diffraction edge. They were spaced at intervals of 1 m. Two methods were examined to detect the reference signal: a microphone located nearby the blower and a vibration pickup (velocity meter) attached to the blower (Fig. 14). In both cases, the coherence function between the outputs of the detection sensor and one of the error sensors (located at $z = 0.5$ m) were measured. This coherence function provides a very convenient method for obtaining an estimate of the effect of an active control system as discussed in Ref. 8.

Results are shown in Fig. 15. The maximum coherencies at around 125 Hz were 0.9996 and 0.9984 in the cases of the microphone and the vibration pickup, respectively. This result suggests that the microphone might give a slightly better control effect. Using the microphone as a detecting sensor, however, sometimes introduces another problems, i.e., that of the performance and the stability of the system due to the
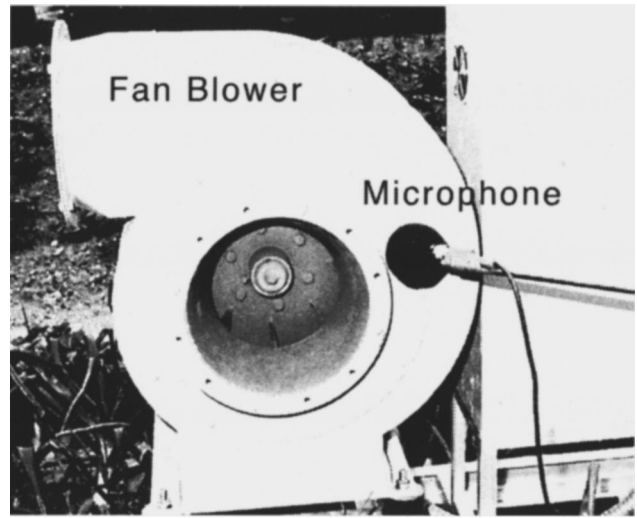
(a)

(b)

FIG. 14. Two methods to detect the reference signal: (a) microphone located nearby the blower; (b) vibration pickup (velocity meter) attached to the blower.

external noise and the feedback path from the secondary sources. Since the measured attenuation in the previous experiment was not so large (about 6 dB), any such problems might wipe out these effectiveness. On the other hand, it is generally difficult for the external noise and the secondary sound to affect the mechanically generated vibration of the noise source. To avoid any disadvantages described above, we thus adopted the vibration pickup as the detecting sensor.

B. Results and discussion

Figures 16–18 show the results. All the levels shown here are 125-Hz third octave-band level. In each figure, (a) and (b) correspond to heights $y = 1$ m and 2 m, respectively, and (c) is the effect of the active control. The three lines in Fig. 16(a) and (b) show the measured level with the active control turned off, that with the active control turned on, and that of the background noise alone.

FIG. 13. Geometry of the barrier, the noise source, the control system, and the receivers.

(a) Microphone as a detecting sensor

(b) Vibration pickup as a detecting sensor
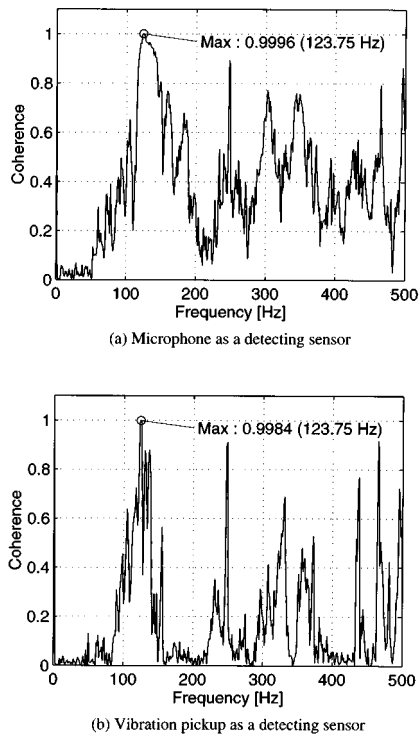
FIG. 15. The coherence function between the outputs of the detection sensor and the error sensor.

In Fig. 16(a) and (b), the sound pressure level with the active system turned off decreased somewhat irregularly as x increases. The same trend can be seen in the case of the active system turned on, but diminished by 1–8 dB. The large attenuation was hardly attained at the dips, i.e., the points where the sound pressure was low with the system turned off.

Figures 17 and 18 show similar results. Also in these cases, the attenuation is low at the dips of the sound pressure level with the system turned off and around 2-dB amplification occurred at one point ($x = 10$ m) in Fig. 18(a). At other points, 1–10 dB attenuation could be achieved. It should be noted that the secondary sound worked as if to "equalize" the irregular distribution of the sound pressure level along each x line.

In this experiment, moderate attenuation could be attained at almost all the receiver points. These results suggests the possibility of the application of the active control to the diffracted sound with realistic noise sources.

Needless to say, there are many problems which should be solved before practical application of our strategy. For instance, the wind speed sometimes reached around 6–7 m/s in this experiment. The external noise due to the wind caused amplification of the sound pressure at some receiver points during the iterative process of the adaptive system. To cope with such problems, more robust algorithms and an experimental study would be needed for each application.

III. CONCLUDING REMARKS

The present study provided experimental verification of our previously developed theory for the active suppression of

(a) y = 1 m   x (m) ———▶

(b) y = 2 m   x (m) ———▶

(c) Effect of the Active Control   x (m) ———▶
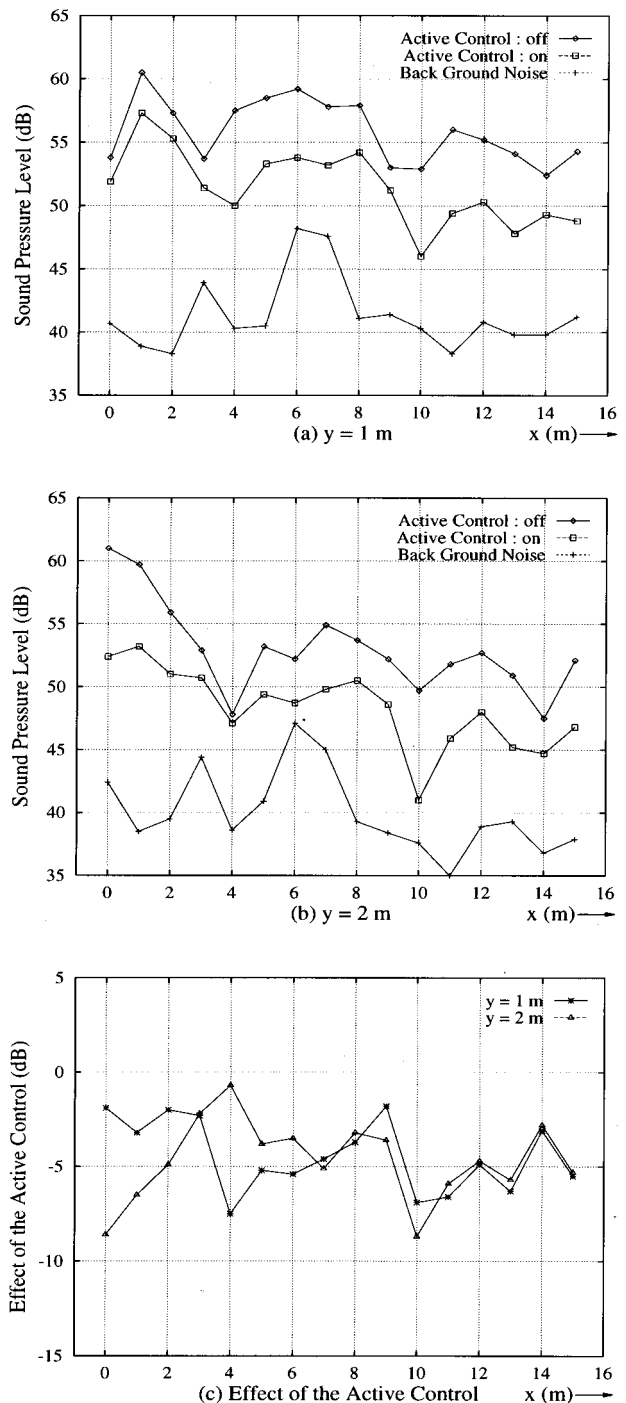
FIG. 16. Sound pressure level and the effect of the active control (along the line $z = 0$ m).

sound diffracted by a noise barrier. In the first experiment, two independent controllers were used, and the sound pressure at a total of ten points on the diffraction edge was minimized by using adaptive signal processing. The results are summarized as follows: (i) the two adaptive systems worked stably, having no detrimental effects on each other; and (ii) the control provided more than 6 dB of attenuation at receivers located about 50 m from the barrier.

The results suggest that additional adaptive systems could be employed to provide more channels for active control. Since 6 dB of attenuation at a receiver 50-m distant
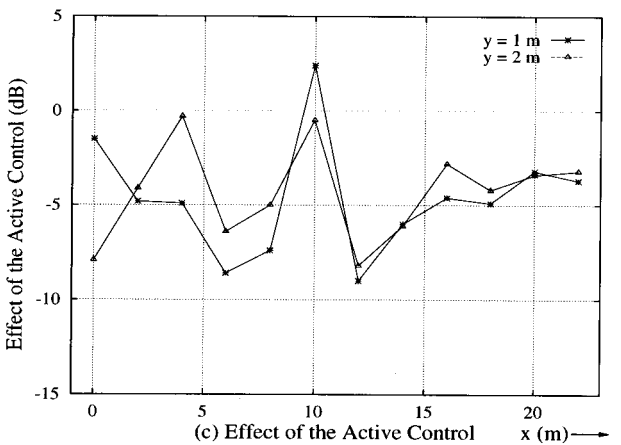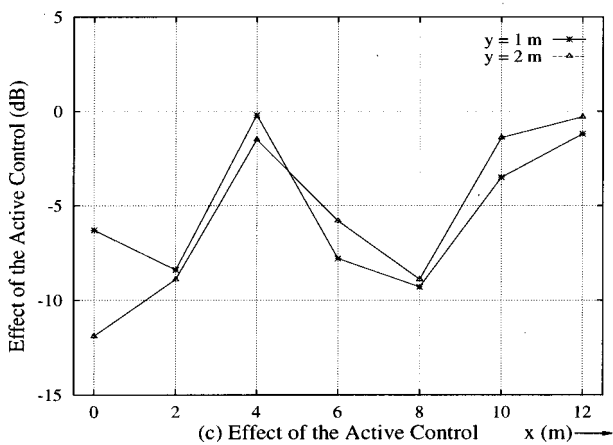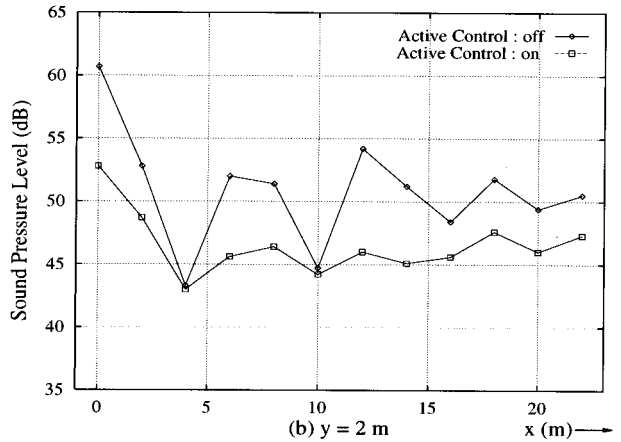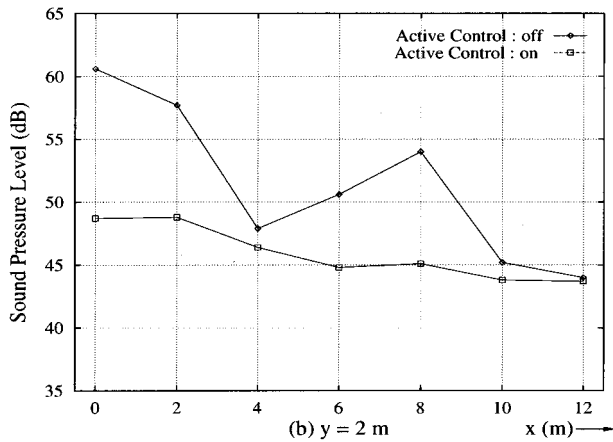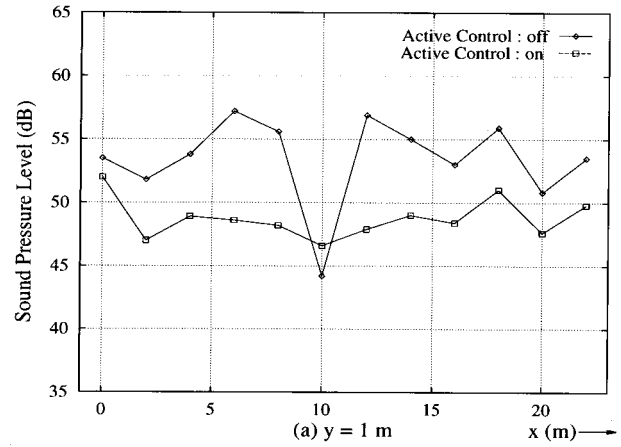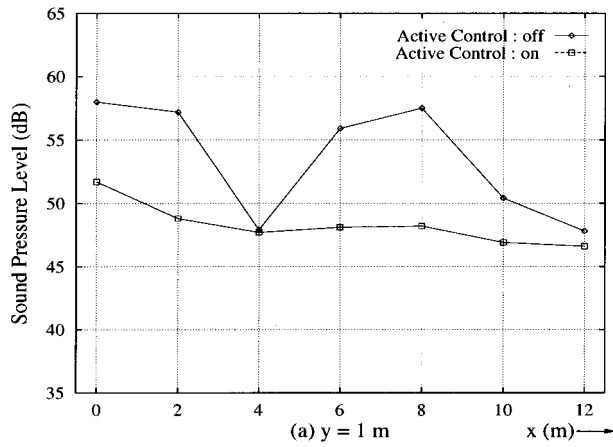
FIG. 17. Sound pressure level and the effect of the active control (along the line $z = -4$ m).

FIG. 18. Sound pressure level and the effect of the active control (along the line $z = 4$ m).

from the barrier is equivalent to increasing the height of the barrier from 2 to 5 m, it is likely that the system has practical application.

In the second experiment, much more realistic noise source, the fan blower, was introduced. Two detection methods of the primary signal were examined and the vibration pick up was used in the actual control. A moderate effect was obtained at almost all the receiver points. From these two experiments, we believe that the validity of the principal of our strategy, i.e., cancelling the sound pressure at the diffraction edge, was proved experimentally.

There are, of course, some problems to be solved before actual use. In the second experiment, the external noise due to the strong wind adversely affect the adaptive system, and the negative effects were sometimes observed during the iterative process. Our experiment was completed within a few hours; however, if the system is operated continuously, the fluctuation of the transfer responses between the sources and the error sensors (e.g., due to the changes of the temperature) may become an another cause of instability in the system. Introducing a more robust algorithm would be needed to cope with such problems.

In addition, durability of the system would be necessary, especially in the case of outdoor use, and countermeasures to plural and moving primary sources might be fairly difficult problems for application to traffic noise control. These problems, including the development of new adaptive algorithm, are a current subject of investigation.

ACKNOWLEDGMENTS

[1] A. Omoto and K. Fujiwara, ''Active canceling of edge potential of noise barrier,'' J. Acoust. Soc. Jpn. **47**, 801–808 (1991) (in Japanese).

[2] A. Omoto and K. Fujiwara, ''A study of an actively controlled noise barrier,'' J. Acoust. Soc. Am. **94**, 2173–2180 (1993).

[3] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes* (North Holland, Amsterdam, 1987), p. 329.

[4] S. J. Elliott, C. C. Boucher, and P. A. Nelson, ''The effect of acoustic coupling on the stability and convergence of independent controllers,'' in *Proceedings of Inter Noise '91* (Australian Acoustical Society, Sydney, 1991), pp. 157–160.

[5] S. J. Elliott and C. C. Boucher, ''Interaction between Multiple Feedforward Active Control Systems,'' IEEE Trans. Acoust. Speech Audio Process. **2**, 521–530 (1994).

[6] 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. **ASSP-35**, 1423–1434 (1987).

[7] H. Hamada, ''Signal processing for active control,'' in *Proceedings of the 1991 Acoustical Society of Japan International Symposium on Active Control of Sound and Vibration* (Acoustical Society of Japan, Tokyo, 1991), pp. 33–44.

[8] P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, New York, 1992).

Calculation of the dynamic air flow resistivity of fiber materials

Viggo Tarnow

Department of Applied Engineering and Production, Technical University of Denmark, Bygning 358,
DK 2800 Lyngby, Denmark

The acoustic attenuation of acoustic fiber materials is mainly determined by the dynamic resistivity to an oscillating air flow. The dynamic resistance is calculated for a model with geometry close to the geometry of real fiber material. The model consists of parallel cylinders placed randomly. Two cases are treated: flow perpendicular to the cylinder axes, and flow parallel to the axes. In each case two new approximate procedures were used. In the first procedure, one solves the equation of flow in a Voronoi cell around the fiber, and averages over the distribution of the Voronoi cells. The second procedure is an extension to oscillating air flow of the Brinkman self-consistent procedure for dc flow. The procedures are valid for volume concentration of cylinders less than 0.1. The calculations show that for the density of fibers of interest for acoustic fiber materials the simple self-consistent procedure gives the same results as the more complicated procedure based on average over Voronoi cells. Graphs of the dynamic resistivity versus frequency are given for fiber densities and diameters typical for acoustic fiber materials. © *1997 Acoustical Society of America.* [S0001-4966(97)01209-5]

PACS numbers: 43.58.Bh, 43.20.Hq, 43.20.Jr, 43.55.Ev [SLE]

LIST OF SYMBOLS

A	vector velocity potential
A_z	Z component of vector potential
a	radius of a cylinder
$a_1, a_2, a_3 \ldots$	integration constants
a_t	radius of tubes in theory of porous materials
b	square root of area per cylinder
$b_1, b_2, b_3 \ldots$	integration constants
$c_1, c_2, c_3 \ldots$	integration constants
$d_1, d_2, d_3 \ldots$	integration constants
$F_\parallel$	force on a cylinder per length for parallel flow
$F_\perp$	force on a cylinder per length for perpendicular flow
g, h	integration constants
i	0 or 1
k	wave vector defined in Eq. (10)
$k'_\parallel$	wave vector in iterative computation for parallel flow
$k'_\perp$	wave vector in iterative computation for perpendicular flow
P	complex power
$P(\rho)$	complex power as a function of the cell radius
p	local air pressure
$\langle p \rangle$	mean pressure=sound pressure
$p_b(S)$	probability density of area of Voronoi polygons
$R_\parallel$	resistivity for parallel flow
$R'_\parallel$	resistivity due to forces on cylinders for parallel flow
$R_\perp$	resistivity for perpendicular flow
$R'_\perp$	resistivity due to forces on cylinders for perpendicular flow
$R(\omega)$	scalar resistivity
$\mathbf{R}(\omega)$	resistivity tensor
r	radius. $r = \sqrt{x^2 + y^2}$
S	area of Voronoi polygons
$\bar{S}$	mean area of Voronoi polygons
t	time
u	air particle velocity vector
$\langle u \rangle$	mean velocity=sound velocity
u_r, u_θ	components of air velocity in polar coordinate system
u_z	component of velocity in Z direction
V	macroscopic volume
w	complex variable
x, y, z	rectangular coordinates
$Z_i(w)$	special Bessel function—solution of Bessel's equation—Eq. (9)
η	viscosity of air
θ	angle in polar coordinate system; $\tan \theta = y/x$
ρ	radius of a circle with area of Voronoi polygon
ρ_0	mass density of air in equilibrium
$\rho(\omega)$	effective dynamic mass density of air
σ	dc resistivity
ϕ	scalar velocity potential
Ψ	volume velocity
Ω	porosity Eq. (16)
ω	cyclic frequency

INTRODUCTION

We assume that sound waves in a fiber material can be described by a linear theory, and we use the complex notation for monochromatic waves with the time factor $e^{-i\omega t}$, where ω is the cyclic frequency and t the time. The resistivity tensor for fiber materials $\mathbf{R}(\omega)$ gives the gradient of the air pressure when the air flow velocity is given; the relation being

$$-\langle \boldsymbol{\nabla} p \rangle = \mathbf{R}(\omega) \langle \mathbf{u} \rangle, \tag{1}$$

where $\langle p \rangle$ is the mean pressure and $\langle \mathbf{u} \rangle$ the mean particle velocity vector, and mean values are taken over the volume of the fiber material. The air is regarded as an incompressible fluid in this definition of the resistivity. The resistivity will be computed from the volume concentration of the fibers and the mean diameter of the fibers, with the purpose of being able to calculate the acoustic properties such as sound velocity and attenuation in fibrous materials. These properties also depend on the compressibility, which has been calculated in Tarnow.[1] In the present paper it is assumed that the fibers do not move, because it is apparently not necessary to take this into account in order to calculate the velocity and attenuation of audible sound propagation in glass wool at normal densities.[2]

The standard method of describing sound propagation in fibrous materials is the theory of porous materials, which models the medium as a solid with air filled tubes of an equal radius much smaller than the wavelength of interest. The medium is regarded as isotropic. Therefore the dynamic resistivity is a scalar $R(\omega)$. Some authors use the effective mass density of air $\rho(\omega)$, which is related to the dynamic resistivity by $\rho(\omega) = -i\omega R(\omega)$. The effective mass density was computed for a circular tube by Zwikker and Kosten,[3] who gave the formula

$$\rho(\omega) = \frac{\rho_0}{1 - 2J_1(ka_t)/ka_t J_0(ka_t)}, \tag{2}$$

where ρ_0 is the static mass density of air, a_t is the radius of the tubes, $J_0(w)$ and $J_1(w)$ are the usual Bessel functions, and

$$k = \sqrt{i\omega\rho_0/\eta}, \tag{3}$$

where η is the viscosity of air.

In porous materials the effective mass density is somewhat higher than predicted by Eq. (2) because a part of the cross section is filled with solid material, which makes the air velocity in the tubes higher than the mean velocity. This depends on the density of fibers, but for acoustic fiber materials the volume density of fibers is in most cases smaller than 2%, and for the sake of simplicity we therefore neglect a small increase in resistivity due to this. Therefore Eq. (2) gives the resistivity of the fiber material.

Biot[4] used the same formula written in a different way, and the author considered the coupling of waves in the solid matrix to waves in the fluid, which fills the pores. Lambert[5] studied open-cell foam and used the same formula, and the author suggested ways to choose a reasonable value of the tube radius. Lambert and Tesar[6] applied the formula to fibrous material and found a way to determine the tube radius from experimental data for the static resistivity. Attenborough[7] also used Eq. (2) for one tube, but for the bulk medium ka_t in Eq. (2) was replaced by a quantity containing the resistivity, the porosity, the tortuosity, and a pore shape factor. Stinson[8] computed the dynamic mass density for rectangular tubes.

Stinson and Champoux[9] showed that the pore shape factor suggested by Attenborough must depend on frequency; they also suggest a new formula for computing the dynamic mass density from the dc flow resistivity, tortuosity, porosity,

and a new pore shape factor. Allard and Champoux[10] gave a formula for the dynamic mass density that only requires knowledge of the dc flow resistivity.

It is the goal of the present work to calculate the dynamic resistivity of fibrous materials from the microscopic geometry of the material without introducing adjustable parameters that are used to fit predictions from models to experimental data. The purpose is to obtain a better physical understanding of the dynamic resistivity of fiber materials in order to be able to predict it when the microscopic geometry is known. The diameters of the fibers can be found by microscopy and the volume density from mass density.

We consider models that have a geometry similar to the microscopic geometry of fiber materials. In the models it is assumed that the fibers all have the same diameter, are parallel, and randomly placed. Flows parallel and perpendicular to fibers are treated.

I. GENERAL THEORY

The Navier–Stokes equation for incompressible stationary flow can be found in Sec. 15 of Landau and Lifschitz.[11] We assume the air velocity is small—low Reynold's number. In this case the equation for the velocity vector is

$$\rho_0 \frac{\partial \mathbf{u}}{\partial t} = -\boldsymbol{\nabla} p + \eta \nabla^2 \mathbf{u}, \tag{4}$$

where $\mathbf{u}$ is the air particle velocity vector, p is the local air pressure, and t the time. In Pierce[12] the viscosity of air is given, $\eta = 1.846 \times 10^{-5}$ kg/(ms) at the temperature 27 °C, its temperature dependence can also be found in Pierce.[12]

The resistivity will be used for computing the acoustic properties of fiber materials such as glass wool. A continuum model is used to describe sound propagation in the fiber materials. The only material parameters are compressibility and resistivity. The compressibility has been computed in Tarnow.[1] In a propagating sound wave at a fixed time there are places where the compression is zero. In such places one needs the resistivity of incompressible flow in order to compute the energy loss and the kinetic energy. Therefore, the assumption of only one resistivity requires that it is calculated for incompressible flow. Thus one must neglect the compressibility of air when one calculates the resistivity.

The incompressibility of the flow is expressed by

$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0. \tag{5}$$

The boundary conditions for Eqs. (4) and (5) are that the velocity on the cylinders walls equals zero. The procedure for computing the resistivity is to solve these equations when
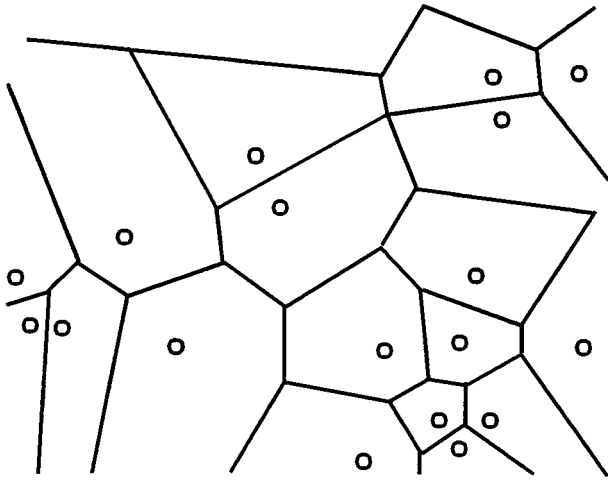
FIG. 1. The Voronoi partition of the plane. The points shown as circles are placed randomly over the plane. The lines are constructed by drawing lines from each point to the nearest neighbor point and drawing lines that bisects the lines between points. Only the bisecting lines are shown.

the mean gradient of the air pressure is given, where the mean is taken over the volume of the fiber material. From the velocity calculated from these equations one computes the mean velocity and finds the resistivity from Eq. (1).

We assume the cylinders are randomly placed. There seems to not be experimental information about the distribution of the fibers in real materials. One could assume that the fibers were placed in a regular lattice, but this is certainly not the case in glass wool. Measurements to be reported later show that for a given density the resistivity calculated for the regular lattice is too high. The resistivity for dc. flow in a regular lattice was calculated in Tarnow.[13] Random placement seems to be the best model, and it can explain the measurement of sound velocity and attenuation that will be reported later.

A cell method is used to find the microscopic velocity. We take the parallel cylinders as given and cut them with a perpendicular plane; in this plane each of the centers of the cylinders is surrounded by a so-called Voronoi polygon. It is constructed by drawing lines from each center to the nearest neighbor centers and drawing lines that bisects the lines between centers. These lines form the Voronoi polygons (Fig. 1). The area S of the Voronoi polygons has a probability $p_b(S)dS$ in the interval S, $S+dS$

$$p_b(S) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \left(\frac{S}{\overline{S}}\right)^{\alpha-1} \exp\left(-\alpha\,\frac{S}{\overline{S}}\right) \frac{1}{\overline{S}}, \quad (6)$$

where $\alpha = 3.61$, $\Gamma(\alpha)$ is the gamma function, and $\overline{S}$ is the mean value of the area of the Voronoi polygons. The formula was given by Andrade and Fortes.[14]

The procedure is the same in each case; the Navier–Stokes equation is solved approximately in a Voronoi polygon, and the flow through the polygon is computed. The average flow is computed by using Eq. (6). From the average flow in a polygon the average velocity is calculated, and the resistivity is found from this. This procedure has in the case

of dc flow been compared with other methods of computing the resistivity, and it was found that it yielded reliable results.[13]

In order to check the accuracy of the cell approximations, the resistivity was also computed by a self-consistent procedure, which is similar to the calculation of the dc resistivity by Brinkman's approximation.[15]

II. FLOW PARALLEL TO CYLINDERS

The flow is assumed to be parallel to the cylinders; in this case the gradient of the pressure does not depend on the space point,[11] and the Navier–Stokes equation is easier to solve than for flow perpendicular to the cylinders. The air velocity is everywhere in the direction of the flow (see Ref. 11), and we assume it to be along the Z axis of a rectangular X-Y-Z coordinate system. The Navier–Stokes equation becomes

$$-i\omega\rho_0 u_z = -\frac{\partial p}{\partial z} + \eta\left(\frac{\partial^2 u_z}{\partial x^2} + \frac{\partial^2 u_z}{\partial y^2}\right). \quad (7)$$

The cylinders are parallel and randomly placed with constant density over a plane perpendicular to the cylinders, and the velocity must be zero at their surface. In this case the incompressibility condition Eq. (5) is always fulfilled.

A. The cell approximation

The Navier–Stokes equation is first solved with only one cylinder whose axis is on the Z axis of the coordinate system. We use polar coordinates with radius r and angle θ. The boundary conditions are: $u_z(a) = 0$, where a is the radius of the cylinder, and $\partial u_z/\partial r = 0$ on a circle $r = \rho$. The polygons are approximated by circles with different ρ for each Voronoi polygon. We imagine that the velocity function is plotted above the plane of Fig. 1. Because the velocity is zero on the cylinders, the plot has a ridge close to the perimeter of each Voronoi polygon in Fig. 1, and the velocity gradient perpendicular to the sides of the polygon is close to zero. When the polygons are replaced by circles with radius ρ, it is reasonable to assume that the gradient along the radius is zero. In the case of dc flow this was used in Tarnow,[13] where it gave accurate results.

When we use the usual complex notation for harmonic variations, the solution to the Navier–Stokes equation (7) in polar coordinates is

$$u_z(r) = \frac{1}{i\omega\rho_0}\frac{\partial p}{\partial z}\left[1 + \frac{J_1(k\rho)}{Z_0(ka)}H_0^1(kr) - \frac{H_1^1(k\rho)}{Z_0(ka)}J_0(kr)\right], \quad (8)$$

with the usual Bessel functions, and we have defined a special Bessel function $Z_i(w)$ of the complex variable w,

$$Z_i(w) = H_1^1(k\rho)J_i(w) - J_1(k\rho)H_i^1(w) \quad (9)$$

with

$$k = \sqrt{\frac{i\omega\rho_0}{\eta}}. \quad (10)$$

The function in Eq. (9) will used later. The volume velocity through a cell of a radius ρ is

$$\Psi(\rho)=\int_a^\rho u_z(r)2\pi r\,dr. \tag{11}$$

After integration the flux becomes

$$\Psi(\rho)=\frac{1}{i\omega\rho_0}\frac{\partial p}{\partial z}\left[\pi\rho^2-\pi a^2\right.$$
$$\left.-\frac{2\pi a(J_1(k\rho)H_1^1(ka)-H_1^1(k\rho)J_1(ka))}{kZ_0(ka)}\right]. \tag{12}$$

The resistivity can be found by calculating the mean velocity from Eq. (12) and using the definition of resistance Eq. (1). But it is more informative and accurate to compute the resistivity in the following way. Equation (7) is transformed by use of the Gauss integral theorem over the area of the cell; thus

$$-i\omega\rho_0\Psi(\rho)=-(\pi\rho^2-\pi a^2)\frac{\partial p}{\partial z}-\eta 2\pi a\left[\frac{\partial u_z}{\partial r}\right]_{r=a}. \tag{13}$$

The average over the cells is computed, and one gets

$$-i\omega\rho_0\langle\Psi\rangle=-(\pi\langle\rho^2\rangle-\pi a^2)\frac{\partial p}{\partial z}-\eta 2\pi a\left\langle\left[\frac{\partial u_z}{\partial r}\right]_{r=a}\right\rangle, \tag{14}$$

where $\langle\ \rangle$ means average.

We divide both sides of the equation with the flux and rearrange the terms, and get the resistivity $R_\parallel$ from $R_\parallel=-(\partial p/\partial z)/\langle u_z\rangle$ and $\langle u_z\rangle=\langle\Psi\rangle/b^2$, where b^2 is the mean area of the Voronoi cells. The result is

$$R_\parallel=\Omega^{-1}\left[-i\omega\rho_0-\frac{\eta 2\pi a}{\langle\Psi\rangle}\left\langle\left[\frac{\partial u_z}{\partial r}\right]_{r=a}\right\rangle\right], \tag{15}$$

where the porosity Ω is

$$\Omega=\frac{b^2-\pi a^2}{b^2}. \tag{16}$$

The first term in the bracket of Eq. (15) is the resistivity one would find if there were no cylinders. The mean value of the flux through a unit cell is found from Eqs. (12) and (6); thus

$$\langle\Psi\rangle=\int_a^\infty\Psi(\rho)p_b(\pi\rho^2)2\pi\rho\,d\rho. \tag{17}$$

The integral was calculated by numeric integration using Eqs. (12) and (6). By differentiation of Eq. (8) $\partial u_z/\partial r$ was found; then $[\partial u_z/\partial r]_{r=a}$ was substituted for Ψ in Eq. (17), and $\langle[\partial u_z/\partial r]_{r=a}\rangle$ was found by numeric integration. The lower dashed line in Fig. 2 shows the real part of the resistivity, where the parameters used correspond to glass wool with mass density 16 kg/m^3, cylinder diameter 6.8 μm, and $b=76$ μm. The volume density of fibers is 0.0063. The lower dashed line in Fig. 3 shows the real part of the resistivity for cylinder diameter 6.8 μm, and $b=40$ μm, corresponding to a volumen concentration of 0.040.
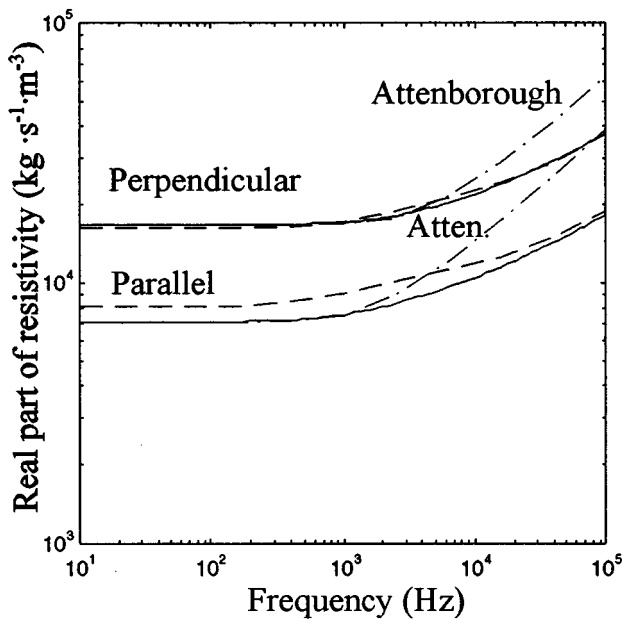
FIG. 2. The real part of the dynamic resistivity. The area per cylinder is b^2 with $b=76$ μm, and the diameter of the cylinders is 6.8 μm. These numbers are valid for glass wool with a density of 16 kg/m^3. The lower dashed line is for flow parallel with cylinders calculated by the cell method. The lower full line is for parallel flow by the Brikman method. The upper dashed line is for flow perpendicular to cylinders by the cell method. The upper full line is for perpendicular flow by the Brinkman method. The dashed-dotted lines are found by the Attenborough approximation; the lower one is for parallel flow, and the upper one for perpendicular flow.
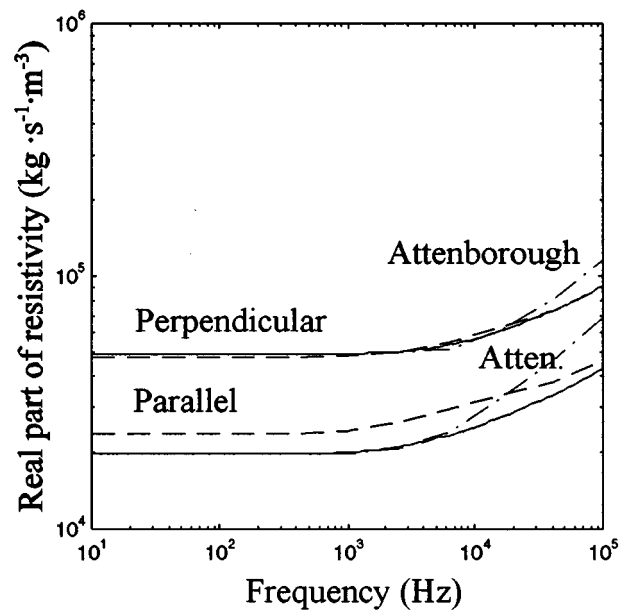
FIG. 3. The real part of the dynamic resistivity. The area per cylinder is b^2 with $b=50$ μm, and the diameter of the cylinders is 6.8 μm. These numbers are valid for glass wool with a density of 40 kg/m^3. The lower dashed line is for flow parallel with cylinders calculated by the cell method. The lower full line is for parallel flow by the Brinkman method. The upper dashed line is for flow perpendicular to cylinders by the cell method. The upper full line is for perpendicular flow by the Brinkman method. The dashed-dotted lines are found by the Attenborough approximation; the lower one is for parallel flow, and the upper one for perpendicular flow.

B. Brinkman's self-consistent approximation

A simple self-consistent (iterative) calculation of Brinkman has been used for many years to calculate the resistivity of dc flow.[15] It can also be used to calculate the dynamic resistivity as will be shown in the following. Only one cylinder is considered, placed at the origin of a coordinate system; the influence of all the other cylinders on the one at the origin is taken care of by introducing a constant body force in the Navier–Stokes equation instead of the discrete forces from the cylinders. The force is proportional to the velocity, and the factor of proportionality is the ''resistivity'' $R'_\parallel$. The equation for the velocity of the air becomes

$$-i\omega\rho_0 u_z = -\frac{\partial p}{\partial z} + \eta\left(\frac{\partial^2 u_z}{\partial x^2} + \frac{\partial^2 u_z}{\partial y^2}\right) - R'_\parallel u_z. \qquad (18)$$

The force on a volume element is the gradient of the pressure; it equals the inertial forces plus the forces from the cylinders. That is, $-\partial p/\partial z = -i\omega\rho_0 u_z + R'_\parallel u_z$. From this and Eq. (1) follows $R_\parallel = R'_\parallel - i\omega\rho_0$.

We regard the velocity as a function of the distance r from the z axis, and the boundary conditions are

$$u_z(a) = 0 \qquad (19)$$

and

$$u_z(\infty) = \langle u \rangle. \qquad (20)$$

A solution to Eq. (18) that fulfils these boundary conditions is

$$u_z = \langle u \rangle \left[1 - \frac{H_0^1(k'_\parallel r)}{H_0^1(k'_\parallel a)} \right], \qquad (21)$$

where

$$k_\parallel'^2 = (i\omega\rho_0 - R'_\parallel)/\eta. \qquad (22)$$

The force on one cylinder per length $F_\parallel$ is

$$F_\parallel = 2\pi a\,\eta\,\frac{\partial u_z}{\partial r}\bigg|_{r=a}. \qquad (23)$$

We require that the distributed forces per volume equal the forces at the origin times the number of cylinders per volume. Thus

$$R'_\parallel \langle u \rangle = F_\parallel b^2. \qquad (24)$$

Set the velocity from Eq. (21) into (23) and use (24) to get

$$R'_\parallel = \frac{\eta 2\pi k'_\parallel a H_1^1(k'_\parallel a)}{b^2 H_0^1(k'_\parallel a)}. \qquad (25)$$

Equations (22) and (25) are solved by an iterative process. First the value of $k'_\parallel = i/b$ is chosen and set into (25); then a value of the resistivity $R'_\parallel$ is computed. From the new value of the resistivity a new value of $k'_\parallel$ is computed by (22), the new value of the wave vector is set into (25), and a new resistivity is computed, etc. After four iterations it was not possible to see any change on a graph of the resistivity as a function of the frequency. The calculation is ''self-consistent'' when the resistivity is constant from one iteration to the next one.

The lower full line of Fig. 2 shows the real part of the resistivity computed for $a = 3.4\ \mu$m, $b = 76\ \mu$m. The lower full line of Fig. 3 shows the real part of the resistivity computed for $a = 3.4\ \mu$m, $b = 50\ \mu$m. The Brinkman method gives nearly the same result as the cell method; this has also been found by calculation of dc flow resistivity,[15] which shows that the Brinkman approximation gives reasonably accurate values of the resistivity for volume concentration <0.1. The cell method tends to give too high a value of the resistivity because it was assumed that the cylinder was placed in the center of the Voronoi cells, which is not the case for a number of Voronoi cells, as one can see in Fig. 1. The Brinkman approximation is quite simple to implement on a computer, and accurate enough for the calculation of the resistivity of fiber materials.

III. FLOW PERPENDICULAR TO CYLINDERS

The flow is now perpendicular to the cylinders, which we assume are all parallel and have the same diameter, they are randomly distributed with constant density. We use a rectangular coordinate system with the Z axis in the direction of the cylinder axes, and the X axis in the direction of the flow. The Navier–Stokes equation (4) shall be solved, but in this case the solution is more complicated. The local air velocity $\mathbf{u}$ is expressed by the scalar potential ϕ and the vector potential $\mathbf{A}$:

$$\mathbf{u} = \nabla\phi + \nabla\times\mathbf{A}. \qquad (26)$$

Here the scalar potential must fulfil the Laplace equation

$$\nabla^2\phi = 0, \qquad (27)$$

because the flow is incompressible, Eq. (5).

An equation for the vector potential can be found by substitution of Eq. (26) into the Navier–Stokes equation (4),

$$-\rho_0 i\omega(\nabla\phi + \nabla\times\mathbf{A}) = \nabla p + \eta\nabla^2(\nabla\phi + \nabla\times\mathbf{A}). \qquad (28)$$

This equation is fulfilled if

$$-\rho_0 i\omega\nabla\phi + \nabla p = 0 \qquad (29)$$

and

$$-\rho_0 i\omega\mathbf{A} + \eta\nabla^2\mathbf{A} = 0. \qquad (30)$$

We use a polar coordinate system with r, θ as coordinates, where

$$x = r\cos\theta, \quad y = r\sin\theta. \qquad (31)$$

The scalar potential is a solution to Laplace's equation (27), and can be written as a Fourier series

$$\phi = a_0 \ln r + \sum_{n=1}^{\infty} [(a_n r + b_n r^{-n})\cos n\theta$$

$$+ (c_n r + d_n r^{-n})\sin n\theta]. \qquad (32)$$

Here a_n, b_n, c_n, d_n are constant. The vector potential lies only in the z direction because the flow is two dimensional. It is a solution to Eq. (30), and can be written as a Fourier series
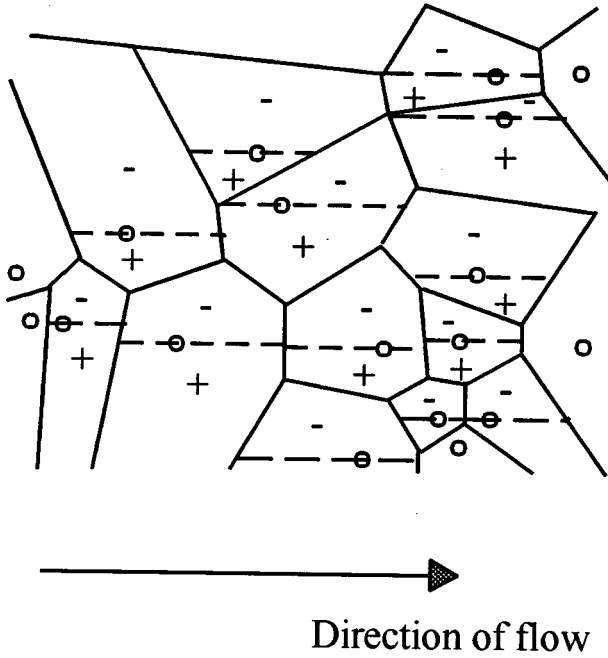
Direction of flow

FIG. 4. The rotation of the velocity in the cells. The dashed lines are parallel to the direction of the mean flow, which is from left to right. In the upper part of each cell the rotation is negative, in the lower part positive. The rotation is zero close to the border between the upper and lower part of neighbor cells.

$$A_z = a_0' H_0^1(kr) + b_0' J_0(kr) + \sum_{n=1}^{\infty} \left[(a_n' H_n^1(kr) + b_n' J_n(kr)) \right.$$

$$\left. \times \cos n\theta + (c_n' H_n^1(kr) + d_n' J_n(kr)) \sin n\theta \right], \tag{33}$$

where k is defined in Eq. (10), and a_n', b_n', c_n', d_n' are constants.

A. The cell approximation

By the flow between the cylinders of radius a the rotation of the velocity vanishes close to the perimeter of each Voronoi cell. In Fig. 4 the flow is from left to right. The rotation $\partial u_y / \partial x - \partial u_x / \partial y$ is negative in the upper part of each cell and positive in the lower part. Where the upper part of one cell meets the lower part of another cell along the border, the rotation tends to be zero. This argument comes from a paper by Kuwabara[16] on dc flow. The same assumption was used in Tarnow[13] for calculation of the dc flow resistivity. It was shown that it gave reasonable results.

We use the cell approximation to compute the resistivity. We first consider a simple problem. One cylinder of a radius a is given; this cylinder is surrounded coaxially with a cylinder surface of a radius ρ. The velocity is symmetric about $\theta = 0$. Therefore, in the scalar potential of Eq. (32), only cosine terms are needed, and due to the circular geometry only the coefficient a_1 and b_1 are different from zero.

The rotation of the velocity is zero for $r = \rho$. This is an approximation to the boundary condition along the perimeter of each Voronoi cell. It will be shown that the rotation of the velocity is

$$\nabla \times \mathbf{u} = \nabla \times \nabla \times \mathbf{A} = -(\rho_0 i \omega / \eta) \mathbf{A}. \tag{34}$$

The vector potential is in the direction of the Z axis perpendicular to the flow and does not depend on the z coordinate because the flow is two dimensional. Therefore $\nabla \cdot \mathbf{A} = 0$, and the vector identity $\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ gives $\nabla \times \nabla \times \mathbf{A} = -\nabla^2 \mathbf{A}$. Equation (34) then follows from Eq. (30). From Eq. (34) one sees that the vanishing of the rotation of the velocity implies that the vector potential must be zero. On the cylinder surface with radius ρ we require that the vector potential vanish.

We first solve the Navier–Stokes equation in the circular cell. In order to make the formulas a little more simple, we use the Bessel function of Eq. (9) $Z_i(kr)$, which for $i = 1$ is zero at the cylinder surface with radius ρ.

The velocity is symmetric about $\theta = 0$, and the velocity equals the rotation of the vector potential; therefore, only the sine terms are kept in Eq. (33). In the circular cell only the first sine term is needed. Thus the vector potential can be written

$$A_z = g Z_1(kr) \sin \theta, \tag{35}$$

where g is an integration constant.

The velocity in the direction of the radius u_r is from Eqs. (26), (32), (35):

$$u_r = \left(a_1 - \frac{b_1}{r^2} + \frac{1}{r} g Z_1(kr) \right) \cos \theta, \tag{36}$$

and the velocity perpendicular to the radius u_θ is

$$u_\theta = -\left(a_1 + \frac{b_1}{r^2} + g k Z_1'(kr) \right) \sin \theta. \tag{37}$$

Here the prime on the Bessel function indicates the derivative. The velocity must be zero at the internal cylinder surface, $r = a$; from this and the above formulas one gets

$$a_1 - \frac{b_1}{a^2} + \frac{1}{a} g Z_1(ka) = 0 \tag{38}$$

and

$$a_1 + \frac{b_1}{a^2} + g k Z_1'(ka) = 0. \tag{39}$$

The pressure follows from Eqs. (29) and (32):

$$p = i \omega \rho_0 \left[a_1 r + \frac{b_1}{r} \right] \cos \theta. \tag{40}$$

For the random medium, we assume that the mean pressure gradient $\langle \nabla p \rangle$ is equal to the pressure drop over the cell divided by the cell diameter 2ρ, and get

$$a_1 + \frac{b_1}{\rho^2} = \frac{\langle \nabla p \rangle}{i \omega \rho_0}. \tag{41}$$

Equations (38), (39), and (41) are three linear equations in the three unknowns a_1, b_1, g, when $\langle \nabla p \rangle$ is known. When they are solved, the velocity and pressure in the cell are known.

The resistivity will be computed by considering the energy flow. In general it is more accurate to use energy meth-

ods to compute resistivity instead of using the definition Eq. (1) directly. The complex power P of the medium in the volume V is

$$P = -\langle \boldsymbol{\nabla} p \rangle^* \cdot \langle \mathbf{u} \rangle V, \tag{42}$$

where $*$ means complex conjugation. The real part of this is the heat generated per volume. From the definition of resistivity Eq. (1) one gets $-\langle \boldsymbol{\nabla} p \rangle = R_\perp \langle \mathbf{u} \rangle$, where $R_\perp$ is the resistivity for perpendicular flow. From this and Eq. (42) one gets

$$R_\perp = \frac{|\langle \boldsymbol{\nabla} p \rangle|^2}{P/V}. \tag{43}$$

The power per length $P(\rho)$ inside the cylinder with radius ρ can be calculated from

$$P(\rho) = - \oint_S p^* \mathbf{u} \cdot d\mathbf{A}, \tag{44}$$

where the integration is done over the cylinder surface with radius ρ. After integration and use of (9) one gets

$$P(\rho) = \frac{(\rho^2/a^2 - 1)Z_0(ka) + (2/ka)Z_1(ka)}{(\rho^2/a^2 + 1)Z_0(ka) - (2/ka)Z_1(ka)} \, \pi\rho^2 \, \frac{|\nabla p|^2}{i\omega\rho_0}. \tag{45}$$

The resistivity is calculated from

$$R_\perp = \frac{|\nabla p|^2 b^2}{\langle P \rangle}, \tag{46}$$

where the angular brackets mean an average value computed by the probability density in Eq. (6) in the following way:

$$\langle P \rangle = \int_a^\infty P(\rho) p_b(\pi\rho^2) 2\pi\rho \, d\rho. \tag{47}$$

Here $P(\rho)$ is the function defined in (45), and $\bar{S} = b^2$ in (6).

B. Brinkman's self-consistent approximation

The dynamic resistivity for flow perpendicular to cylinders can be calculated by a self-consistent procedure similar to the one used for parallel flow. To the Navier–Stokes equation is added a body force $-R'_\perp \mathbf{u}$ that equals the force from the cylinders per volume, where $R'_\perp$ is the resistivity due to the forces on the cylinders. The Navier–Stokes equation becomes

$$-i\omega\rho_0 \mathbf{u} = -\boldsymbol{\nabla} p + \eta \nabla^2 \mathbf{u} - R'_\perp \mathbf{u}. \tag{48}$$

Only one cylinder at the origin of the coordinate system is considered. On the cylinder surface the velocity must be zero,

$$\mathbf{u}(a) = 0; \tag{49}$$

and infinitely far from the cylinder at the origin the velocity equals the mean velocity $\langle u \rangle$, which is in the X direction,

$$\mathbf{u}(\infty) = (\langle u \rangle, 0, 0). \tag{50}$$

A solution to Eq. (48) that fulfils the boundary conditions (49) and (50) can be written in the same way as Eqs. (36) and (37), but the special Bessel function $Z_1(kr)$ is replaced by a

Hankel function, which goes to zero as the radius goes to infinity. One gets

$$u_r = \left(\langle u \rangle - \frac{b_1}{r^2} + \frac{1}{r} \, h H_1^1(k'_\perp r) \right) \cos\theta \tag{51}$$

and

$$u_\theta = - \left(\langle u \rangle + \frac{b_1}{r^2} + h k'_\perp H_1^{1'}(k'_\perp r) \right) \sin\theta. \tag{52}$$

The prime on the Hankel function indicates the derivative, and h is an integration constant. In Eqs. (50) and (51), $k'_\perp$ is given by the right-hand side of Eq. (22) with $R'_\perp$ instead of $R'_\parallel$. The boundary condition Eq. (49), and Eqs. (51) and (52) give by use of the identity $H_1^{1'}(w) = H_0^1(w) - (1/w)H_1^1(w)$,

$$h = - \frac{2\langle u \rangle}{k'_\perp H_0^1(k'_\perp a)}. \tag{53}$$

From Eqs. (48), (49), (51), and (52) one finds

$$b_1 = a^2 \left(1 - \frac{2H_1^1(k'_\perp a)}{k'_\perp a H_0^1(k'_\perp a)} \right) \langle u \rangle. \tag{54}$$

The force on the cylinder at the origin of the coordinate system may now be found from the pressure and the stress tensor. The calculation is similar to the calculation in section 20 of Landau and Lifschitz,[11] where the Stokes formula for the viscous drag on a sphere is computed. There one finds the force written as an integral over the surface of the sphere; and this can be adapted to an integral over a cylinder, which is what is needed here. Section 15 of Landau and Lifschitz[11] also gives the stress tensor in polar coordinates. One finds the force $F_\perp$ per length of the cylinder

$$F_\perp = 2\pi\eta k'_\perp a \left(-k'_\perp a + \frac{2H_1^1(k'_\perp a)}{H_0^1(k'_\perp a)} \right). \tag{55}$$

We require that the body force per volume equals the force from the cylinders per volume similar to Eq. (24), and get the resistivity

$$R'_\perp = \frac{2\pi\eta k'_\perp a}{b^2} \left(-k'_\perp a + \frac{2H_1^1(k'_\perp a)}{H_0^1(k'_\perp a)} \right). \tag{56}$$

We proceed in the same way for the case of parallel flow by iteration. We first choose a value of the wave vector, for example $k'_\perp = i/b$, and calculate a resistivity from (56), then a new wave vector is calculated by Eq. (22), etc. until a stable value of the resistivity is found, which takes about four iterations. The inertia of the air between the cylinders is not included in the resistivity Eq. (56). Therefore the resistivity defined in Eq. (1) is

$$R_\perp = -i\omega\rho_0 + R'_\perp. \tag{57}$$

The upper full line of Fig. 2 shows the real part of the resistivity computed for $a = 3.4 \, \mu$m, $b = 76 \, \mu$m. The upper full line of Fig. 3 shows the real part of the resistivity computed for $a = 3.4 \, \mu$m, $b = 50 \, \mu$m. The Brinkman method gives nearly the same result as the cell method; this has also been found by calculation of dc flow resistivity,[15] which shows that the Brinkman approximation gives reasonably accurate

values of the resistivity for volume concentration <0.1.

IV. DISCUSSION

The assumption of parallel fibers is valid for glass wool plates produced by Glasuld, Scan-Gobain Denmark, which have fibers that are mainly parallel with the largest surface and with one edge, according to information from the producer. The measurements reported in Tarnow[2] were made with sound propagation perpendicular to the largest surface of the plates. In this case one can assume that the sound velocity is perpendicular to the fibers. For other materials and uses one must make and average of the resistance tensor, but this depends on information about the orientation of the fibers in a given material and the direction of sound propagation. This information is not available. But I hope to make new measurements that can give more information on this. Anisotropy of glass wool has been measured by Allard et al.[17] They found that the dc flow resistivity perpendicular to the fibers was two times the resistivity parallel with the fibers. This is in accordance with the low-frequency value found in the present paper and the values calculated in Tarnow.[13] One can show that in the low-frequency limit, $R'_\perp \approx 2R'_\parallel$ follows from Eqs. (25) and (56).

The resistivity has been compared with the model of Attenborough. Because the concentration of fibers is small, one must assume that the porosity and tortuosity is 1. The pore shape ratio must be 0.5 in order to obtain the proper low-frequency limit. We follow the definitions used by Stinson and Champoux.[9] The low-frequency resistivity was calculated by the Brinkman method of the present paper and this was used as the dc resistivity σ. The Attenborough formula for the dynamic effective mass density becomes in this case

$$\rho(\omega) = \frac{\rho_0}{1 - 2J_1(w)/wJ_0(w)}, \qquad (58)$$

with $w = \sqrt{8\rho_0 i\omega/\sigma}$. The result of the calculation is shown in Figs. 2 and 3 as the dashed-dotted line. In Fig. 2 there are differences between the Brinkman approximation and the Attenborough formula at high audio frequencies. The Attenborough model assumes a circular tubes, but tubes with other cross-section shapes will not give very different high-frequency dependence because tubes have greater surfaces than fibers.

The imaginary part of the resistivity is approximately proportional to the frequency. In order to show graphically the difference between the different methods of calculating, the real part of the dynamic effective mass density is given. This is related to the imaginary part of the resistivity by $\text{Im}\{R\} = -\omega \cdot \text{Re}\{\rho(\omega)\}$ which follows from $\rho(\omega) = -i\omega R(\omega)$. Figure 5 shows $\text{Re}\{\rho(\omega)\}/\rho_0$ for $a = 3.4 \ \mu\text{m}$, $b = 76 \ \mu\text{m}$. In Fig. 5 the full line is for flow parallel with cylinders (fibers) calculated by the Brinkman method Eqs. (22) and (25) and $R_\parallel = R'_\parallel - i\omega\rho_0$. The dashed line is for parallel flow calculated by the Attenborough formula Eq. (58). The dotted line is for perpendicular flow; it was calculated by Brinkman's method Eqs. (22) and (56) and $R_\perp = R'_\perp - i\omega\rho_0$. The dashed-dotted line is for perpendicular
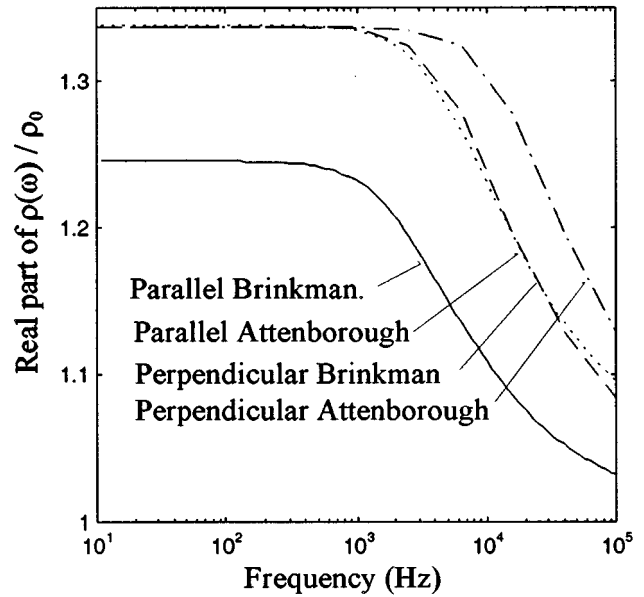
FIG. 5. The real part of the effective mass density. This is related to the imaginary part of the resistivity by $\text{Im}\{R\} = -\omega \ \text{Re}\{\rho(\omega)\}$. The figure shows $\text{Re}\{\rho(\omega)\}/\rho_0$, where ρ_0 is the mass density of air. The area per cylinder is b^2 with $b = 76 \ \mu\text{m}$, and the diameter of the cylinders is 6.8 μm corresponding to a mass density of glass wool of 16 kg/m³. The full line is for flow parallel with cylinders (fibers) calculated by the Brinkman method. The dashed line is for parallel flow; it was calculated by the Attenborough formula Eq. (58). The dotted line is for perpendicular flow; it was calculated by Brinkman's method. The dashed-dotted line is for perpendicular flow; it was calculated by Attenborough's Eq. (58).

flow by Attenborough's Eq. (58). It is obvious that there are differences between the results of the different methods of calculation.

V. CONCLUSION

The dynamic flow resistivity tensor of fiber materials has been calculated for a model that consists of parallel cylinders of equal diameter. The resistivity was calculated from the density of cylinders and their diameter. Two cases were considered: flow parallel with and flow perpendicular to cylinders. From this and the symmetry of the resistivity tensor all the elements of the tensor are known.

The calculations were done with two approximations: a cell approximation and Brinkman's approximation. They gave results close to each other, but the Brinkman procedure is simpler to implement on a personal computer; it requires only a few lines of program in a high level mathematical program such as MATLAB. The accuracy of Brinkman's procedure is estimated to be of the order of 10% for fiber densities normally used in materials for acoustics; this was based on the calculations in this paper and Tarnow[13] and Sangani and Yao.[15]

[1] V. Tarnow, ''Compressibility of air in fibrous materials,'' J. Acoust. Soc. Am. **99**, 3010–3017 (1996).

[2] V. Tarnow, ''Measurement of sound propagation in glass wool,'' J. Acoust. Soc. Am. **97**, 2272–2281 (1995).

[3] C. Zwikker and C. W. Kosten, *Sound Absorbing Materials* (Elsevier, Amsterdam, 1949).

[4] 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).

[5] R. F. Lambert, ''The acoustical structure of highly porous open-cell foam,'' J. Acoust. Soc. Am. **72**, 879–887 (1982).

[6] R. F. Lambert and J. S. Tesar, ''Acoustic structure and propagation in highly porous, layered, fibrous materials,'' J. Acoust. Soc. Am. **76**, 1231–1237 (1984).

[7] K. Attenborough, ''Acoustical characteristics of rigid fibrous adsorbents,'' J. Acoust. Soc. Am. **73**, 785–799 (1983).

[8] M. R. Stinson, ''The propagation of plane sound waves in narrow and wide circular tubes, and generalization to uniform tubes of arbitrary cross-sectional shape,'' J. Acoust. Soc. Am. **89**, 550–558 (1991).

[9] M. R. Stinson and Y. Champoux, ''Propagation of sound and the assignment of shape factors in models of porous materials having simple pore geometries,'' J. Acoust. Soc. Am. **91**, 685–695 (1992).

[10] J.-F. Allard and Y. Champoux, ''New empirical equations for sound propagation in rigid frame fibrous materials,'' J. Acoust. Soc. Am. **91**, 3346–3353 (1992).

[11] L. D. Landau and E. M. Lifschitz, *Fluid Mechanics* (Pergamon, Oxford, 1996).

[12] A. D. Pierce, *Acoustics* (Acoustical Society of America, New York, 1989), p. 513.

[13] V. Tarnow, ''Air flow resistivity of fibrous acoustic materials,'' J. Acoust. Soc. Am. **100**, 3706–3713 (1996).

[14] P. N. Andrade and M. A. Fortes, ''Distribution of cell volumes in a Voronoi partition,'' Philos. Mag. B **58**, 671–674 (1988).

[15] A. S. Sangani and C. Yao, ''Transport processes in random arrays of cylinders. II. Viscous flow,'' Phys. Fluids **31**, 2435–2444 (1988).

[16] S. Kuwabara, ''The forces experienced by randomly distributed parallel circular cylinders or spheres in a viscous flow at small Reynolds numbers,'' J. Phys. Soc. Jpn. **14**, 527–532 (1959).

[17] J.-F. Allard, R. Boudier, and A. l'Esperance, ''Anisotropy effect in glass wool on normal impedance in oblique incidence,'' J. Sound Vib. **114**, 233–238 (1987).

Detection of narrow-band signals in spherically invariant noise

Luciano Izzo[a] and Mario Tanda

Università di Napoli Federico II, Dipartimento di Ingegneria Elettronica, Via Claudio, 21 I-80125 Napoli, Italy

This paper deals with noncoherent discrete-time detection of a narrow-band signal subject to slow and nonselective fading and embedded in correlated non-Gaussian noise modeled as a spherically invariant random process whose modulating random variable is continuous. At first, an asymptotic sufficient statistic for an arbitrary fading law is derived; then, the asymptotically optimum detector for Rayleigh-distributed amplitude fluctuations is synthesized. The detection structure implementation requires the knowledge, but for a scale factor, of the correlation function of the noise, but is independent of the distribution function of the modulating random variable. The performance of the asymptotically optimum detector synthesized for Rayleigh fading is assessed via computer simulations. The results show that the performance degradation with respect to the fully optimum performance is scarcely significant, even for low values of the sample size. Moreover, in highly non-Gaussian noise the proposed detector largely outperforms the fully optimum detector synthesized under the correlated Gaussian noise assumption. © *1997 Acoustical Society of America.* [S0001-4966(97)04608-0]

PACS numbers: 43.60.Cg, 43.60.Gk [JLK]

INTRODUCTION

The fully optimum (in the Neyman–Pearson sense) algorithms for signal detection in non-Gaussian noise environments are generally very difficult to implement physically. Therefore, to obtain canonical detection structures of easy implementation, the assumptions of weak signal and noise samples modeled as independent and identically distributed random variables (RVs) have usually been adopted (see Refs. 1–11). Locally optimum detection structures, which maximize the slope of the power function at zero signal level while keeping a fixed false-alarm rate, have been derived. Moreover, since in weak-signal detection one uses a relatively large sample size to obtain a reasonable value for the detection probability, asymptotically optimum (AO) detectors have been synthesized, whose operating characteristics approach optimality as the sample size increases and the signal amplitude vanishes (Refs. 9, 10, and 12).

In many cases of practical interest, the assumption of independent noise samples is inappropriate, which implies that the detectors synthesized on this assumption do not retain their optimality properties.

A very attractive non-Gaussian noise model for the development of detection algorithms is correlated noise assumes that the disturbance can be viewed as a Gaussian process modulated by a non-negative process slowly varying compared to the Gaussian one. This model, first proposed by Hall in Ref. 13, is physically consistent with some noise phenomena of great interest, such as atmospheric noise (Ref. 13), sea clutter (Ref. 14), and under-ice and shallow-water noise (Ref. 15). Moreover, its validity is confirmed by several experimental results (see Refs. 13, 15, and 16).

In most cases of practical interest, both theoretical considerations and experimental results have shown that the co-herence time of the modulating process is larger than the observation time (Refs. 13–15, 17, and 18). Consequently, the non-Gaussian noise can be modeled in the observation time as a spherically invariant random process (SIRP), which is completely characterized by the knowledge of the univariate probability density function (PDF) of the modulating RV, the mean, and the covariance function (Ref. 19). Such a model is a generalization of the Gaussian model and, hence, retains most of its attractive properties. Furthermore, commonly adopted families of univariate PDFs, such as the generalized Laplace, the generalized Gamma, the generalized Gaussian with a shape parameter less than or equal to two, and the generalized Cauchy are admissible as univariate PDFs of SIRPs (see Refs. 19 and 20).

The optimum detection of a signal in spherically invariant (SI) noise whose modulating RV is a strictly positive discrete RV has been considered in Ref. 17 for both continuous-time and discrete-time observations. Moreover, it has been shown (Ref. 21) that the classical linear (coherent case) and envelope (noncoherent case) receivers optimized against Gaussian noise retain their optimality properties when adopted for reception of equally likely and equal-energy signals embedded in SI noise. Finally, the locally optimum array detector (Ref. 20) and a constant false-alarm rate detector (Ref. 22) for a random signal in SI noise have been synthesized.

The present paper deals with the synthesis of the AO detector for the noncoherent discrete-time detection of a narrow-band signal subject to slow and nonselective fading and embedded in correlated SI noise whose modulating RV is continuous. A typical application is in active sonar, where the noise is suitably modeled as correlated and non-Gaussian and, moreover, the narrow-band transmitted signal is affected by amplitude fluctuations, which are usually modeled as Rayleigh distributed RVs (Ref. 7). In the paper, at first an

---

[a]Electronic mail: izzo@nadis.dis.unina.it

asymptotic sufficient statistic for an arbitrary fading law is derived; then, the AO detector for the interesting case of Rayleigh fading is synthesized.

The performance of the AO detector synthesized for Rayleigh fading, in terms of detection probability for a fixed false-alarm rate, is assessed via computer simulations by assuming the generalized Cauchy model for the univariate PDF of the noise. Such a PDF family is quite flexible encompassing a wide class of PDFs and has been frequently adopted to model impulsive noise environments (see Refs. 2, 10, 13, 23, and 24). Moreover, a comparison with the asymptotic performance of the fully optimum detector is made to evaluate the performance degradation of the proposed AO detector when it operates with small and moderately large sample sizes. Finally, the performances of the AO detector are compared with those of the fully optimum detector synthesized under the correlated Gaussian noise assumption.

We emphasize that the results of this paper are consistent with those obtained in Ref. 17 not only for the form of the solution, but also for its general properties, such as the information needed to implement the algorithm and to evaluate its performance. With reference to the case of discrete-time observations, the paper extends the work presented in Ref. 17 providing an explicit solution to the problem of noncoherent detection of a faded narrow-band signal embedded in SI noise with a continuous modulating RV for which noise whitening can be applied.

Finally, we observe that some care must be taken in applying the results obtained in the paper, since in the detector synthesis it is implicitly assumed (see Ref. 10) that the observation time increases as the sample size increases. In fact, the SI noise model assumes that the modulating RV has a constant value over the observation time, and this is a questionable assumption when the observation time becomes arbitrarily long. Note, however, that, for the noise environment considered in the simulations, the proposed AO detector works well even for low values of the sample size.

I. SYNTHESIS OF THE DETECTION STRUCTURE

The detection problem under consideration can be represented by the following hypothesis test:

$$
\begin{aligned}
H_0: & \quad \widetilde{\mathbf{x}} = \widetilde{\mathbf{z}}, \\
H_1: & \quad \widetilde{\mathbf{x}} = \theta A e^{j\phi} \widetilde{\mathbf{v}} + \widetilde{\mathbf{z}},
\end{aligned}
\tag{1}
$$

where θ is a positive parameter (that will be specified in the following), a tilde identifies complex quantities, and $\widetilde{\mathbf{x}}$, $\widetilde{\mathbf{z}}$ and $\widetilde{\mathbf{v}}$ are N-dimensional row vectors whose entries are samples drawn from the complex envelopes of the received signal, the noise, and the narrow-band signal to be detected. The random variable A, which assumes non-negative values, takes into account the presence of a slow (in comparison with the observation time) amplitude fading. The RV ϕ is assumed to be uniformly distributed over a 2π interval (noncoherent detection). The RVs A and ϕ and the noise vector $\widetilde{\mathbf{z}}$ are mutually independent.

The SI noise vector $\widetilde{\mathbf{z}} = \mathbf{z}_c + j\mathbf{z}_s$, according to Yao's representation theorem (Ref. 19), can be written as

$$
\widetilde{\mathbf{z}} = a \widetilde{\mathbf{n}}_g,
\tag{2}
$$

where a is a non-negative modulating RV independent of $\widetilde{\mathbf{n}}_g = \mathbf{n}_{gc} + j\mathbf{n}_{gs}$, which is a zero-mean Gaussian vector characterized by a $2N \times 2N$ correlation matrix $\sigma_g^2 \mathbf{K}$ with σ_g^2 the common variance of the inphase and quadrature components.

The matrix $\mathbf{K}$ admits the Cholesky decomposition $\mathbf{K} = \mathbf{C}\mathbf{C}^T$, where $\mathbf{C}$ is a $2N \times 2N$ invertible lower triangular matrix and the superscript T denotes transpose operation. Therefore, the correlated noise vector $\widetilde{\mathbf{z}}$ can be transformed into a white noise vector $\widetilde{\mathbf{w}} = \mathbf{w}_c + j\mathbf{w}_s$ by the relationship $(\mathbf{w}_c, \mathbf{w}_s) = (\mathbf{z}_c, \mathbf{z}_s)(\mathbf{C}^{-1})^T$. Moreover, the closure property of the SI vectors under deterministic linear transformations (see Ref. 19) assures that the vector $\widetilde{\mathbf{w}}$ is still SI with the same modulating RV a as $\widetilde{\mathbf{z}}$, zero mean, and correlation matrix $\mathbf{K}_{\widetilde{\mathbf{w}}} = E(a^2)\sigma_g^2 \mathbf{I}$, where $E(a^2)$ is the mean-square value of the RV a and $\mathbf{I}$ is the $2N \times 2N$ identity matrix. Finally, the theorem of reversibility assures that the detector synthesized on the basis of the hypothesis test

$$
\begin{aligned}
H_0: & \quad \widetilde{\mathbf{y}} = \widetilde{\mathbf{w}}, \\
H_1: & \quad \widetilde{\mathbf{y}} = \theta A e^{j\phi} \widetilde{\mathbf{s}} + \widetilde{\mathbf{w}}
\end{aligned}
\tag{3}
$$

retains the optimality properties of the detector synthesized starting from Eqs. (1). In Eqs. (3), $\widetilde{\mathbf{y}} = \mathbf{y}_c + j\mathbf{y}_s$ and $\widetilde{\mathbf{s}} = \mathbf{s}_c + j\mathbf{s}_s$ are obtained by $(\mathbf{y}_c, \mathbf{y}_s) = (\mathbf{x}_c, \mathbf{x}_s)(\mathbf{C}^{-1})^T$ and $(\mathbf{s}_c, \mathbf{s}_s) = (\mathbf{v}_c, \mathbf{v}_s)(\mathbf{C}^{-1})^T$, where $\mathbf{v}_c$ and $\mathbf{v}_s$ are the inphase and quadrature components of $\widetilde{\mathbf{v}}$.

The optimum test for H_0 vs H_1 (in the Neyman–Pearson sense) compares the likelihood ratio

$$
\Lambda(\widetilde{\mathbf{y}}) = E_{A,\phi}[\Lambda(\widetilde{\mathbf{y}}|A,\phi)]
\tag{4}
$$

to some threshold chosen to achieve the desired false-alarm rate. In Eq. (4), $E_{A,\phi}[\cdot]$ denotes the statistical expectation with respect to the RVs A and ϕ, and the conditional likelihood ratio can be written as

$$
\begin{aligned}
& \Lambda(\widetilde{\mathbf{y}}|A,\phi) \\
& = \frac{\displaystyle\int_0^{+\infty} a^{-2N} \exp\left(-\frac{\|\widetilde{\mathbf{y}} - \theta A e^{j\phi}\widetilde{\mathbf{s}}\|^2}{2a^2\sigma_g^2}\right) p_a(a)\, da}{\displaystyle\int_0^{+\infty} a^{-2N} \exp\left(-\frac{\|\widetilde{\mathbf{y}}\|^2}{2a^2\sigma_g^2}\right) p_a(a)\, da},
\end{aligned}
\tag{5}
$$

where $\|\cdot\|$ denotes Euclidean norm and $p_a(\cdot)$ is the univariate PDF of a.

The fully optimum detector for the case where a is a discrete RV is presented in Ref. 17, in which the more general case of random signal detection is considered. In this paper, an AO detection structure for the case where the modulating RV is continuous is synthesized.

To introduce the AO detector, one can consider an approach similar to that followed in Ref. 25. To this end, let us substitute

$$
\frac{\|\widetilde{\mathbf{y}}\|^2}{2a^2\sigma_g^2} = Nx
\tag{6}
$$

in both integrals in (5). It results that

$$\Lambda(\widetilde{\mathbf{y}}|A,\phi) = \frac{\int_0^{+\infty} x^{-1/2}\Psi_N(x)\exp\left[-\frac{A^2 N\theta^2 x\|\widetilde{\mathbf{s}}\|^2}{\|\widetilde{\mathbf{y}}\|^2} + \frac{2AN\theta x}{\|\widetilde{\mathbf{y}}\|^2}\operatorname{Re}(\widetilde{\mathbf{s}}^*\widetilde{\mathbf{y}}^T e^{-j\phi})\right]p_a\left(\sqrt{\frac{\|\widetilde{\mathbf{y}}\|^2}{2N\sigma_g^2 x}}\right)dx}{\int_0^{+\infty} x^{-1/2}\Psi_N(x)p_a\left(\sqrt{\frac{\|\widetilde{\mathbf{y}}\|^2}{2N\sigma_g^2 x}}\right)dx}, \tag{7}$$

where $\operatorname{Re}\{\cdot\}$ and $*$ stand for real part and conjugation (respectively), and the function

$$\Psi_N(x) = \frac{N^N}{\Gamma(N)}x^{N-1}\exp(-Nx) \quad (x \geq 0) \tag{8}$$

has been defined, where $\Gamma(\cdot)$ is the Gamma function. Let us now observe that the function $\Psi_N(x)$ can be viewed as the PDF of a Gamma RV (Ref. 26) with statistical expectation equal to 1 and variance equal to $1/N$. Then, for $N\to\infty$ the variance tends to zero and the RV converges in mean square to its statistical expectation, which implies that

$$\Psi_N(x) \to \delta(x-1) \quad \text{as} \quad N\to\infty, \tag{9}$$

where $\delta(x-1)$ is Dirac's delta function located at $x=1$. Therefore, the following relation holds:

$$\Lambda(\widetilde{\mathbf{y}}|A,\phi) \doteq \exp\left[-\frac{A^2\theta^2\|\widetilde{\mathbf{s}}\|^2}{(1/N)\|\widetilde{\mathbf{y}}\|^2}\right]$$
$$\times \exp\left[\frac{2A\theta}{(1/N)\|\widetilde{\mathbf{y}}\|^2}\operatorname{Re}(\widetilde{\mathbf{s}}^*\widetilde{\mathbf{y}}^T e^{-j\phi})\right], \tag{10}$$

where $\doteq$ indicates that both sides of (10) converge to one and the same function as $N\to\infty$. Finally, performing the expectation of both sides of (10) with respect to ϕ and accounting for (4), one obtains the following decision statistic for the AO detector:

$$T^{\mathrm{AO}}(\widetilde{\mathbf{y}}) = E_A\left[\exp\left(-\frac{A^2\theta^2\|\widetilde{\mathbf{s}}\|^2}{(1/N)\|\widetilde{\mathbf{y}}\|^2}\right)\right.$$
$$\left.\times I_0\left(\frac{2A\theta}{(1/N)\|\widetilde{\mathbf{y}}\|^2}|\widetilde{\mathbf{s}}^*\widetilde{\mathbf{y}}^T|\right)\right], \tag{11}$$

where $E_A[\cdot]$ denotes the statistical expectation with respect to A and $I_0(\cdot)$ is the modified Bessel function of the first kind and zero order.

The block diagram of the AO detector for the hypothesis test (1) is shown in Fig. 1, where the observation vector
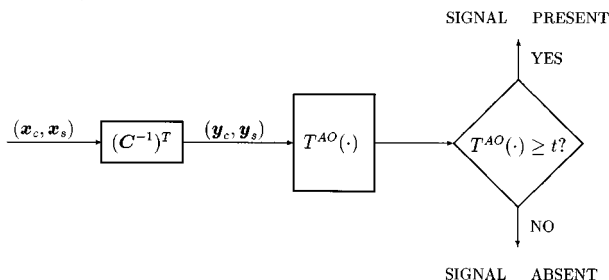
FIG. 1. Asymptotically optimum detection structure.

$(\mathbf{x}_c,\mathbf{x}_s)$ is linearly transformed into the vector $(\mathbf{y}_c,\mathbf{y}_s)$ and then the transformed observations are processed to form the statistic $T^{AO}(\widetilde{\mathbf{y}})$ given by Eq. (11).

The main advantage of the proposed AO detector is that its structure does not depend on the univariate PDF of the modulating RV a. The synthesis of the detection structure, however, requires a priori knowledge, but for a scale factor, of the noise correlation matrix and the knowledge of the fading PDF.

If one assumes that the fading RV A is Rayleigh distributed, from Eq. (11) it follows that the AO decision statistic is given by

$$T_{\mathrm{Ray}}^{\mathrm{AO}}(\widetilde{\mathbf{y}}) = \frac{\theta^2 E(A^2)}{(1/N)\|\widetilde{\mathbf{y}}\|^2[(1/N)\|\widetilde{\mathbf{y}}\|^2 + \theta^2 E(A^2)\|\widetilde{\mathbf{s}}\|^2]}|\widetilde{\mathbf{s}}^*\widetilde{\mathbf{y}}^T|^2$$
$$-\ln\left[1 + \frac{\theta^2 E(A^2)\|\widetilde{\mathbf{s}}\|^2}{(1/N)\|\widetilde{\mathbf{y}}\|^2}\right], \tag{12}$$

where $E(A^2)$ is the mean-square value of A.

As is shown in the Appendix A, in white SI noise both the detection probability and the false-alarm rate of the considered AO detectors depend on the signal $\widetilde{\mathbf{s}} \equiv \widetilde{\mathbf{v}}$ only through the mean (over fading) signal-to-noise ratio

$$\mathrm{SNR}_0 = \frac{\theta^2 E(A^2)\|\widetilde{\mathbf{s}}\|^2}{2E(a^2)\sigma_g^2}, \tag{13}$$

resulting so unaffected by the shape of the signal to be detected. Consequently, to avoid singular detection problems, that is, to assure that SNR_0 is finite and not zero for any value of the sample size N, the parameter θ is assumed to be equal to $\gamma/\sqrt{N}$ with γ a positive constant (see Refs. 9, 10, and 12). On this assumption, $\|\widetilde{\mathbf{y}}\|^2/N$ represents, under both hypotheses and for a given value of a, an asymptotically $(N\to\infty)$ unbiased and consistent estimate of the power $2a^2\sigma_g^2$ of each component of the noise vector $\widetilde{\mathbf{w}}$ conditioned to the modulating RV a. In the light of this observation, it is easy to understand how the transition from Gaussian to SI noise takes place in the detection algorithm. Specifically, it can be seen that the right-hand side of Eq. (11) converges to the likelihood ratio for Gaussian noise (see, e.g., Ref. 27) where, however, the actual SI noise power is introduced.

Finally, since in the particular case of Gaussian noise $p_a(a) = \delta(a-1)$, it results that, for $\theta = \gamma/\sqrt{N}$, $\|\widetilde{\mathbf{y}}\|^2/N$ tends to $2\sigma_g^2$ in mean square and, then, from Eq. (11) it follows that, independently of the fading law, the decision statistic of the AO detector is asymptotically coincident with the decision statistic $|\widetilde{\mathbf{s}}^*\widetilde{\mathbf{y}}^T|$, which is fully optimum in Gaussian noise (Ref. 27).
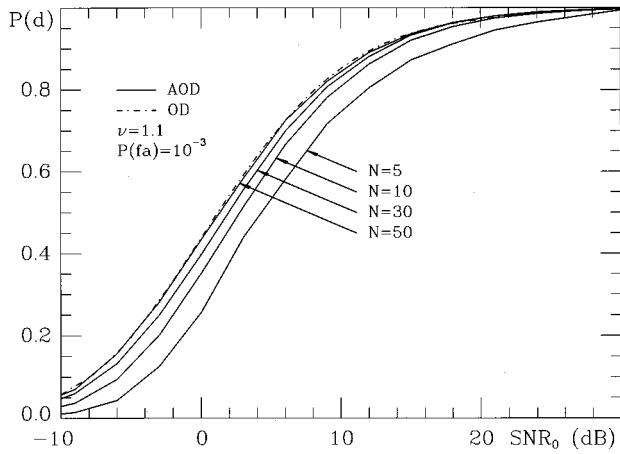
FIG. 2. Detection probability of AOD as a function of the mean signal-to-noise ratio for some values of the sample size N and generalized Cauchy noise with shape parameter $\nu = 1.1$.

FIG. 3. Detection probability of AOD as a function of the mean signal-to-noise ratio for some values of the sample size N and generalized Cauchy noise with shape parameter $\nu = 10$.

II. PERFORMANCE ASSESSMENT

The present section is aimed at assessing the performance of the AO detector synthesized with reference to Rayleigh fading, i.e., the one based on Eq. (12). Moreover, the performance is compared with that of the fully optimum detector in Gaussian noise.

Since the analytical evaluation of the PDFs of the AO decision variable under both hypotheses is an intractable problem, the performance [detection probability $P(d)$ for a fixed false-alarm rate $P(fa)$] has been carried out by Monte Carlo computer simulations. Moreover, the performance has been assessed with reference to the case of white SI noise, where, as is shown in the Appendix A, it depends on the signal to be detected only through the mean signal-to-noise ratio (SNR). In fact, due to the closure property of SIRPs under deterministic linear transformations, the performance in correlated SI noise can be assessed by exploiting the relationship between the mean SNRs at the input and output of the whitening filter.

In the simulations, we have assumed that the amplitude fading is Rayleigh distributed and the white noise vector $\widetilde{\mathbf{w}}$ is characterized by the generalized Cauchy model for the univariate PDF of both quadrature components:

$$p_{wc}(x) = p_{ws}(x) = \frac{[2\sigma_g^2 \nu]^\nu \Gamma(\nu + \frac{1}{2})}{\sqrt{\pi}\Gamma(\nu)[2\sigma_g^2 \nu + x^2]^{\nu + 1/2}}, \qquad (14)$$

where ν is a positive shape parameter. Note that, as ν approaches infinity, the resulting PDF approaches the Gaussian density. Moreover, the variance σ^2 of both quadrature components of the noise is given by $\sigma^2 = \sigma_g^2 \nu/(\nu-1)$ for $\nu > 1$ and is not finite for any value of ν belonging to the interval $(0,1]$, which has not been considered in the simulation experiments.

Figures 2 and 3 present the detection probability $P(d)$ of the AO detector, say AOD, as a function of the mean signal-to-noise ratio SNR_0, for a fixed value ($\nu = 1.1$ in Fig. 2 and $\nu = 10$ in Fig. 3) of the noise shape parameter ν and for some values of the sample size N. The false-alarm rate $P(fa)$ has been fixed at $P(fa) = 10^{-3}$ and the number of
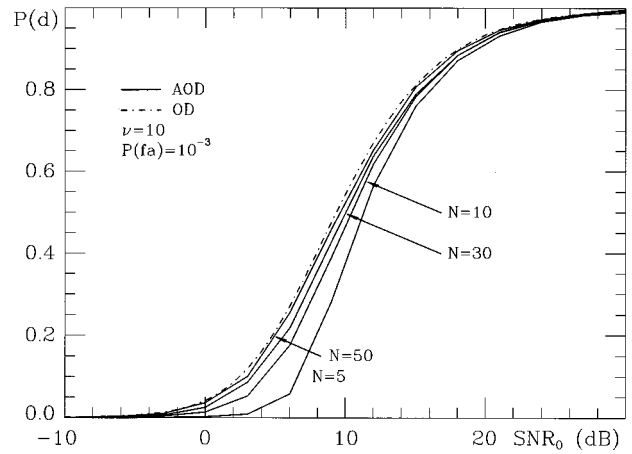
simulation runs per datum has been assumed (as well as in the following experiments) to be $100/P(fa)$. Moreover, in both figures a curve (labeled as OD) referring to the asymptotic ($N \rightarrow \infty$) $P(d)$ is reported, which is just the detection probability of the fully optimum detector in SI noise. Such a probability can be obtained by numerically evaluating the integral

$$P^{\mathrm{OD}}(d) = \int_0^{+\infty} P^{\mathrm{OD}}(d|a) p_a(a) da, \qquad (15)$$

where

$$p_a(a) = \frac{2\nu^\nu}{a^{2\nu+1}\Gamma(\nu)} \exp(-\nu/a^2) \quad (a \geq 0) \qquad (16)$$

is the PDF of the modulating RV a in the case of generalized Cauchy noise (see Refs. 20 and 22) and $P^{\mathrm{OD}}(d|a)$ is the conditional detection probability, which has been evaluated analytically in the Appendix B. The results show that in both noise environments ($\nu = 1.1$ and $\nu = 10$) the performance degradation of the AOD with respect to the optimum performance is scarcely significant even for low values of the sample size N.

The detection probability $P(d)$ of both the AOD and the optimum detector for Gaussian noise, say OGD, is reported in Fig. 4 as a function of SNR_0 and for some values of the noise shape parameter ν. A sample size $N = 30$ and $P(fa) = 10^{-3}$ are assumed. The curves referring to OGD (which are unaffected by the value of N) have been obtained by numerically performing the mean (over a) of the conditional detection probability $P^{\mathrm{OGD}}(d|a)$, which has been evaluated analytically in the Appendix B. The results show that the larger the deviation from the Gaussian distribution ($\nu = \infty$) becomes, the better the signal detectability via the synthesized AOD structure. Moreover, for highly non-Gaussian noise (low values of ν) AOD largely outperforms OGD, whereas, as ν increases, the performances of both detectors tend to be comparable. Finally, in the particular case of
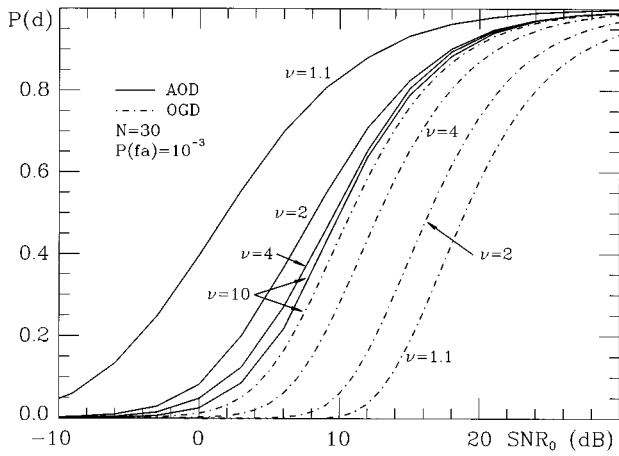
FIG. 4. Detection probability of AOD and OGD as a function of the mean signal-to-noise ratio for some values of the shape parameter ν of the generalized Cauchy noise.

Gaussian noise (curves for $\nu = \infty$), OGD slightly outperforms AOD, since in such a noise environment OGD is a fully optimum detector.

Finally, we underline that the curves presented in the previous figures, referring to the white noise case, can also be utilized to assess the performance in a correlated noise environment provided that, instead of SNR_0, a new value SNR'_0 is considered, which is related (in decibels) to SNR_0 by $SNR'_0 = SNR_0 + \Delta$, where Δ is the SNR gain (or loss) in decibels due to the whitening transformation. For example, let us assume that the signal samples are $\tilde{v}_i = \exp(j2\pi\mu i)$ ($i = 1, 2, \ldots, N$) and the correlation matrix of the noise vector $\tilde{\mathbf{z}}$ has elements

$$K_{\tilde{z}}(i,m) = \begin{cases} \sigma^2 r^{|i-m|}, & i,m = 1,2,\ldots,N \\ & i,m = N+1,\ldots,2N, \\ 0, & \text{otherwise}, \end{cases} \quad (17)$$

where r is the correlation coefficient between adjacent noise samples. The resulting Δ is plotted in Fig. 5 as a function of r, for some values of the signal frequency μ and for $N = 30$. The results can be easily interpreted accounting for the
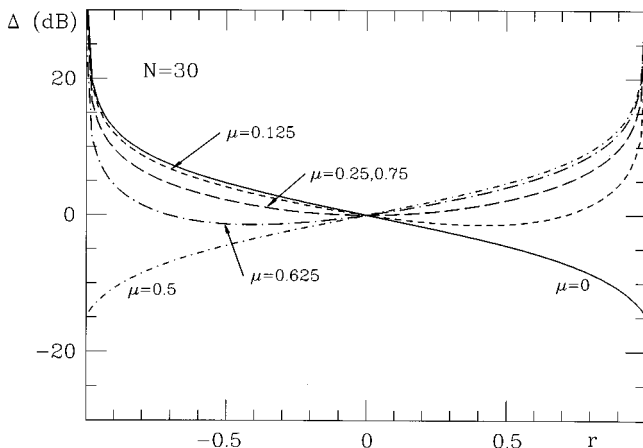
FIG. 5. SNR gain due to the whitening transformation as a function of the noise correlation coefficient r with the signal frequency μ as a parameter.

functional dependence on r of the noise power spectral density. For example, for $\mu = 0.5$ (high-pass signal) and nonnegative values of r (low-pass noise), the amount of overlapping between signal and noise spectra decreases as r increases, leading so to a gain Δ increasing with r.

III. CONCLUSIONS

The synthesis of the asymptotically optimum detector for the noncoherent discrete-time detection of a narrow-band signal subject to slow and nonselective fading and embedded in correlated non-Gaussian noise modeled as a spherically invariant random process with a continuous modulating RV has been presented.

At first, an asymptotic sufficient detection statistic for an arbitrary fading law has been derived; then, the AO detector for the special case of Rayleigh fading has been synthesized. The implementation of the proposed detectors requires *a priori* knowledge, but for a scale factor, of the correlation function of the noise, but is independent of the distribution function of the modulating RV.

The detection probability and the false-alarm rate of both considered detectors in white spherically invariant noise depend on the signal to be detected only through the mean (over fading) SNR, resulting so unaffected by the signal shape. Consequently, due to the closure property of the spherically invariant random processes under deterministic linear transformations, the performance in correlated noise can be easily assessed by exploiting the relationship between the mean SNRs at the input and output of the whitening filter.

The performance (detection probability for a fixed false alarm rate) of the AO detector synthesized for Rayleigh fading has been assessed via Monte Carlo computer simulations by assuming the generalized Cauchy model for the univariate PDF of the noise. The results show that the performance degradation of the AO detector with respect to the fully optimum performance is scarcely significant even for low values of the sample size. Moreover, for highly non-Gaussian noise (i.e., for low values of the noise shape parameter ν), the AO detector largely outperforms the fully optimum detector for correlated Gaussian noise. For large values of ν, the performances of both detectors are comparable, according to the behavior of the generalized Cauchy PDF, which asymptotically ($\nu \to \infty$) approaches the Gaussian PDF.

APPENDIX A

In this Appendix, we show that, under both hypotheses and for any value of N, the PDF of the decision RV of the AO detector (11) [and, consequently, also that of the detector (12)] operating in white spherically invariant noise depends on the signal $\mathbf{s}$ only through the mean (over fading) signal-to-noise ratio, remaining so unaffected by the signal shape.

To simplify the proof of our statement, we observe that it is sufficient to show that both the conditional PDFs

$p_T(\cdot|H_i,A,a)$ $(i=0,1)$ of the decision variable T depend on the signal $\mathbf{s}$ only through the conditional SNR

$$\rho \triangleq \frac{A^2\theta^2\|\widetilde{\mathbf{s}}\|^2}{2a^2\sigma_g^2}. \tag{A1}$$

In fact, on this assumption, it results that

$$p_T(x|H_i) = \int_0^{+\infty}\int_0^{+\infty} p_T(x|H_i,\rho)p_a(a)p_A(A)da\,dA$$

$$= \int_0^{+\infty}\int_0^{+\infty} p_T\left(x\,\middle|\,H_i,\frac{A_0^2}{a_0^2}\,\mathrm{SNR}_0\right)p_{a_0}(a_0)$$

$$\times p_{A_0}(A_0)da_0\,dA_0, \quad i=0,1, \tag{A2}$$

where SNR_0 is the mean SNR and a_0 and A_0 are normalized versions of a and A, respectively, such that their mean-square values are unity.

Following a reasoning similar to that considered in Ref. 28, to carry out our proof, we write the argument of the exponential and Bessel functions involved in (11) as

$$\frac{A^2N\theta^2\|\widetilde{\mathbf{s}}\|^2}{\|\widetilde{\mathbf{y}}\|^2} = \frac{2Na^2\sigma_g^2\rho}{\|\widetilde{\mathbf{y}}_0\|^2+|\widetilde{y}_s|^2} \triangleq Y, \tag{A3}$$

$$\frac{2AN\theta}{\|\widetilde{\mathbf{y}}\|^2}\,|\widetilde{\mathbf{s}}^*\widetilde{\mathbf{y}}^T| = Y\,\frac{\sqrt{2}}{a\sigma_g\sqrt{\rho}}\,|\widetilde{y}_s| \triangleq YZ, \tag{A4}$$

where $\widetilde{\mathbf{y}}$ has been decomposed into its component along $\mathbf{s}$, say $\widetilde{\mathbf{y}}_s = \widetilde{y}_s\mathbf{s}/\|\mathbf{s}\|$, and into the orthogonal component $\widetilde{\mathbf{y}}_0$. At this point, conditioning upon ϕ and $\|\widetilde{\mathbf{y}}_0\|^2 = y'$ leads to

$$H_0: \begin{cases} Y = 2N\left[\dfrac{y'}{a^2\sigma_g^2\rho} + \dfrac{|\widetilde{w}_s|^2}{a^2\sigma_g^2\rho}\right]^{-1}, \\[4mm] Z = \sqrt{2}\,\dfrac{|\widetilde{w}_s|}{a\sigma_g\sqrt{\rho}}, \end{cases} \tag{A5}$$

$$H_1: \begin{cases} Y = N\left[\dfrac{y'}{2a^2\sigma_g^2\rho} + \left|\dfrac{\widetilde{w}_s}{a\sigma_g\sqrt{2\rho}} + e^{j\phi}\right|^2\right]^{-1}, \\[4mm] Z = 2\left|\dfrac{\widetilde{w}_s}{a\sigma_g\sqrt{2\rho}} + e^{j\phi}\right|, \end{cases} \tag{A6}$$

where $\widetilde{\mathbf{w}}_s = \widetilde{w}_s\mathbf{s}/\|\mathbf{s}\|$ is the component of the noise along the signal. Note that the orthogonal vectors $\widetilde{\mathbf{y}}_0$ and $\widetilde{\mathbf{w}}_s$ are mutually independent since the noise is conditionally (upon a) Gaussian. Moreover, in each of the hypotheses, the RVs Y and Z are both functions of one and the same RV whose PDF depends on the signal $\mathbf{s}$ only through ρ. Consequently, accounting for the circular symmetry of the noise, which allows one to absorb ϕ in the uniformly distributed phase of $\widetilde{w}_s$, one obtains that, under both hypotheses, the joint PDF of Y and Z, and hence the conditional PDF $p_T(\cdot|H_i,A,a)$, depends on the signal $\mathbf{s}$ only through ρ, which completes the proof.

APPENDIX B

In this Appendix, we evaluate the asymptotic $(N\to\infty)$ conditional detection probability $P^{\mathrm{OD}}(d|a)$ and false-alarm rate $P^{\mathrm{OD}}(fa|a)$ of the fully optimum detector in white SI noise and Rayleigh fading.

Since for $N\to\infty$ the AO detector performs as well as the fully optimum detector, the conditional probabilities can be written as

$$P^{\mathrm{OD}}(d|a) = \int_\tau^\infty p_{T_0^\infty}(\xi|H_1,a)d\xi \tag{B1}$$

and

$$P^{\mathrm{OD}}(fa|a) = \int_\tau^\infty p_{T_0^\infty}(\xi|H_0,a)d\xi, \tag{B2}$$

where τ is the test threshold and $p_{T_0^\infty}(\cdot|H_i,a)$ $(i=0,1)$ are the asymptotic conditional PDFs of the RV T_0 obtained by replacing in (12) (with $\theta = \gamma/\sqrt{N}$) $\|\widetilde{\mathbf{y}}\|^2/N$ with $2a^2\sigma_g^2$ (see the last observation in Sec. I) and $\|\widetilde{\mathbf{s}}\|^2/N$ with its limiting value P. In other words,

$$T_0 = hX - k, \tag{B3}$$

where

$$h \triangleq \frac{\gamma^2 E(A^2)}{2a^2\sigma_g^2[2a^2\sigma_g^2 + \gamma^2 E(A^2)P]}, \tag{B4}$$

$$k \triangleq \ln\left[1 + \frac{\gamma^2 E(A^2)P}{2a^2\sigma_g^2}\right], \tag{B5}$$

and

$$X \triangleq \frac{1}{N}\,|\widetilde{\mathbf{s}}^*\widetilde{\mathbf{y}}^T|^2. \tag{B6}$$

Since the RV in the magnitude sign in (B6) is conditionally (upon a) Gaussian under both hypotheses H_0 and H_1, the asymptotic PDFs of X are given by

$$p_{X^\infty}(\xi|H_i,a) = \frac{1}{\beta_i}\exp(-\xi/\beta_i) \quad (\xi \geqslant 0), \quad i=0,1, \tag{B7}$$

where

$$\beta_0 \triangleq 2a^2\sigma_g^2 P \tag{B8}$$

and

$$\beta_1 \triangleq \gamma^2 E(A^2)P^2 + 2a^2\sigma_g^2 P. \tag{B9}$$

Therefore, from (B1), (B2), and (B7), it results that

$$P^{\mathrm{OD}}(d|a) = \int_{(\tau+k)/h}^\infty p_{X^\infty}(\xi|H_1,a)d\xi$$

$$= \exp\left\{-\frac{\tau + \ln[1+\lambda(a)]}{\lambda(a)}\right\} \tag{B10}$$

and

$$P^{\text{OD}}(fa|a) = \int_{(\tau+k)/h}^{\infty} p_{X^{\infty}}(\xi|H_0,a)d\xi$$

$$= \exp\left\{-\frac{\{\tau+\ln[1+\lambda(a)]\}[1+\lambda(a)]}{\lambda(a)}\right\}, \tag{B11}$$

where

$$\lambda(a) \triangleq \frac{\gamma^2 E(A^2)P}{2a^2\sigma_g^2} \tag{B12}$$

is the conditional SNR.

Finally, let us note that, for any value of N, the decision statistic of the OGD, that is, the fully optimum detector synthesized under the assumption of Gaussian noise, is equivalent to the statistic X defined in (B6). Then, the conditional probabilities $P^{\text{OGD}}(d|a)$ and $P^{\text{OGD}}(fa|a)$ can be easily obtained taking into account that

$$\int_{\tau}^{\infty} p_X(\xi|H_i,a)d\xi = \exp(-\tau/\alpha_i) \quad (i=0,1), \tag{B13}$$

where

$$\alpha_0 \triangleq 2a^2\sigma_g^2\|\widetilde{\mathbf{s}}\|^2/N \tag{B14}$$

and

$$\alpha_1 \triangleq \frac{\|\widetilde{\mathbf{s}}\|^2}{N}\left[\gamma^2 E(A^2)\frac{\|\widetilde{\mathbf{s}}\|^2}{N} + 2a^2\sigma_g^2\right]. \tag{B15}$$

[1] D. Middleton, ''Canonically optimum threshold detection,'' IEEE Trans. Inf. Theory **IT-12**, 230–243 (1966).

[2] J. H. Miller and J. B. Thomas, ''Detectors for discrete-time signals in non-Gaussian noise,'' IEEE Trans. Inf. Theory **IT-18**, 241–250 (1972).

[3] A. D. Spaulding and D. Middleton, ''Optimum reception in an impulsive interference environment. Part I: Coherent detection,'' IEEE Trans. Commun. **COM-25**, 910–923 (1977).

[4] A. D. Spaulding and D. Middleton, ''Optimum reception in an impulsive interference environment. Part II: Incoherent reception,'' IEEE Trans. Commun. **COM-25**, 924–934 (1977).

[5] H. V. Poor and J. B. Thomas, ''Locally optimum detection of discrete-time stochastic signals in non-Gaussian noise,'' J. Acoust. Soc. Am. **63**, 75–80 (1978).

[6] J. J. Sheehy, ''Optimum detection of signals in non-Gaussian noise,'' J. Acoust. Soc. Am. **63**, 81–90 (1978).

[7] J. Gil Shin and S. A. Kassam, ''Robust detector for narrow-band signals in non-Gaussian noise,'' J. Acoust. Soc. Am. **74**, 527–533 (1983).

[8] G. Fedele, L. Izzo, and L. Paura, ''Optimum and suboptimum space-diversity detection of weak signals in non-Gaussian noise,'' IEEE Trans. Commun. **COM-32**, 990–997 (1984).

[9] L. Izzo and L. Paura, ''Asymptotically optimum space-diversity detection in non-Gaussian noise,'' IEEE Trans. Commun. **COM-34**, 97–103 (1986).

[10] S. A. Kassam, *Signal Detection in Non-Gaussian Noise* (Springer-Verlag, New York, 1988).

[11] L. Izzo, L. Paura, and M. Tanda, ''Signal interception in non-Gaussian noise,'' IEEE Trans. Commun. **40**, 1030–1037 (1992).

[12] B. Levin, *Fondements théoriques de la radiotechnique statistique* (MIR, Moscow, 1976), Vol. III.

[13] H. Hall, ''A new model for 'impulsive' noise phenomena: Application to atmospheric-noise communication channels,'' Tech. Rep. No. 3412-8 and No. 7050-7, Stanford University, Stanford, CA, Apr. 1966.

[14] J. W. Wright, ''A new model for sea clutter,'' IEEE Trans. Antennas Propag. **AP-16**, 217–223 (1968).

[15] C. R. Baker, ''On non-Gaussian signal detection and channel capacity,'' in *Topics in Non-Gaussian Signal Processing*, edited by E. J. Wegman, S. C. Schwartz, and J. B. Thomas (Springer-Verlag, New York, 1989).

[16] K. D. Ward, C. J. Baker, and S. Watts, ''Maritime surveillance radar. Part 1: Radar scattering from the ocean surface,'' Proc. IEE (Pt. F) **137**, 51–62 (1990).

[17] C. R. Baker and A. F. Gualtierotti, ''Likelihood ratios and signal detection for non-Gaussian processes,'' in *Stochastic Processes in Underwater Acoustics, Lecture Notes in Control and Information Sciences*, edited by C. R. Baker (Springer-Verlag, New York, 1986), pp. 154–180.

[18] S. Watts, C. J. Baker, and K. D. Ward, ''Maritime surveillance radar. Part 2: Detection performance prediction in sea clutter,'' Proc. IEE (Pt. F) **137**, 63–72 (1990).

[19] K. Yao, ''A representation theorem and its applications to spherically invariant random processes,'' IEEE Trans. Inf. Theory **IT-19**, 600–608 (1973).

[20] L. Izzo and M. Tanda, ''Array detection of random signals in spherically invariant noise,'' J. Acoust. Soc. Am. **94**, 2682–2690 (1993).

[21] E. Conte, M. Di Bisceglie, and M. Lops, ''Optimum detection of fading signals in impulsive noise,'' IEEE Trans. Commun. **43**, 869–876 (1995).

[22] L. Izzo and M. Tanda, ''Constant false-alarm rate array detection of random signals in spherically invariant noise,'' J. Acoust. Soc. Am. **98**, 931–937 (1995).

[23] R. L. Spooner, ''On the detection of a known signal in a non-Gaussian noise process,'' J. Acoust. Soc. Am. **44**, 141–147 (1968).

[24] S. L. Adams and L. W. Nolte, ''Optimum array detection in fluctuating ambient noise fields,'' J. Acoust. Soc. Am. **58**, 670–678 (1975).

[25] E. Conte, M. Lops, and G. Ricci, ''Distribution-free radar detection in compound-Gaussian clutter,'' in Proceedings of the International Conference ''RADAR '92,'' Brighton, Great Britain, Oct. 1992, pp. 98–101.

[26] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1971), Vol. II.

[27] R. O. Nielsen, *Sonar Signal Processing* (Artech House, London, 1991).

[28] E. Conte, M. Longo, M. Lops, and S. L. Ullo, ''Radar detection of signals with unknown parameters in K-distributed clutter,'' Proc. IEE (Pt. F) **138**, 131–138 (1991).

Spectra and fourth-order cumulant spectra of broadband beamformed data

Roger F. Dwyer

43 South Cobblers Court, Niantic, Connecticut 06357

Sensor arrays play a major role in underwater acoustic applications. They allow propagating signals to be spatially separated after beamforming giving direction of arrival information. They are usually designed by sensor spacing selection for a particular frequency but used over a band of frequencies. For wideband applications, arrays designed with nonuniform sensor spacing have been utilized. But many of the existing acoustic arrays are designed with uniform sensor spacing. When these arrays are used beyond their design frequency spatial resolution is reduced at low frequencies and grating lobes appear at higher frequencies based on the second-order spectrum. To define these effects a theorem and a corollary are stated and proven under general conditions. In contrast, the fourth-order cumulant spectrum provides more flexibility in choosing the frequency response at the output of a beamformer so deliterious effects of wide frequency bandwidths can be reduced. To define these improvements another theorem and corollary are stated and proven for the fourth-order cumulant spectrum. To demonstrate these theoretical results returns from an active sonar system have been simulated. The transmitted waveform is a linear frequency modulated signal and the echo is returned from a spherical target that is either stationary or in motion. The returned signal is beamformed using a one-dimensional line array of uniformly spaced sensors designed for a frequency of 4160 Hz. But the transmitted waveform has a bandwidth from 1 Hz to 8 kHz. The results are given for the matched and mismatched filtered beamformer outputs of a ten-element line array in the frequency and range domains. These results are compared with the fourth-order cumulant spectrum of the beamformer output and its range domain representation. © *1997 Acoustical Society of America.* [S0001-4966(97)03408-5]

PACS numbers: 43.60.Gk [JLK]

INTRODUCTION

Consider a signal $f(t)$ and its corresponding Fourier transform, $F(\omega)$. Often its spectrum, $S(\omega) = F(\omega)F^*(\omega) = |F(\omega)|^2$ plays an important role in analyses. But $S(\omega)$ does not contain phase information. Higher-order generalizations of $S(\omega)$ have been defined that contain phase information.[1–3] Here the fourth-order cumulant spectrum is presented and applied at the output of a beamformer for deterministic data representing acoustic scattering from spheres caused by a broadband active sonar transmitted linear frequency modulated (LFM) waveform. In this new application of the fourth-order cumulant spectrum the phase response will be considered theoretically and the frequency and time domains of a beamformer output when driven well beyond its design frequency will be compared to its corresponding fourth-order cumulant spectrum and its time domain representation. The motivation for this approach is that often it is desirable to be able to extend the frequency range of an array beyond its design frequency without having to redesign and build a new array since this obviously would take time and be costly. On the other hand, when the array output is evaluated at its design frequency, the fourth-order cumulant spectrum still appears to offer performance advantages for the spherical targets. These results are also applicable to the design of microphone arrays.[4,5]

Let $f(t) = \sum_{i=1}^{N_t} g_i(t)$, be composed of a sum of N_t signals, $g_i(t)$, all having a finite duration over the interval, $(-T, T)$, so that the fourth-order cumulant of $f(t)$ is defined as a generalization of a correlation integral,[6]

$$
\begin{aligned}
c_4(\tau_1, \tau_2, \tau_3) = {} & \frac{1}{2T} \int_{-T}^{T} f(t)f(t+\tau_1)f(t+\tau_2)f(t+\tau_3)\,dt \\
& - \frac{1}{2T} \int_{-T}^{T} f(t)f(t+\tau_1)\,dt \, \frac{1}{2T} \int_{-T}^{T} f(t+\tau_2) \\
& \times f(t+\tau_3)\,dt - \frac{1}{2T} \int_{-T}^{T} f(t)f(t+\tau_2) \\
& \times dt \, \frac{1}{2T} \int_{-T}^{T} f(t+\tau_1)f(t+\tau_3)\,dt \\
& - \frac{1}{2T} \int_{-T}^{T} f(t)f(t+\tau_3)\,dt \\
& \times \frac{1}{2T} \int_{-T}^{T} f(t+\tau_1)f(t+\tau_2)\,dt,
\end{aligned}
\tag{1}
$$

where, the first term is the fourth-order moment integral and the other terms are products of autocorrelation integrals. It should be emphasized that this definition of fourth-order cumulant is similar to the fourth-order cumulant discussed by Mendel[7] for random processes and to the fourth-order cumulant discussed by Dwyer[8] for Gaussian amplitude-modulated sinusoids. The fourth-order cumulant reduces to

$$c_4(\tau_1, \tau_2, \tau_3) = \sum_{i_1=1}^{N_t} \sum_{i_2=1}^{N_t} \sum_{i_3=1}^{N_t} \sum_{i_4=1}^{N_t} \left\{ \frac{1}{2T} \int_{-T}^{T} g_{i_1}(t) g_{i_2}(t+\tau_1) g_{i_3}(t+\tau_2) g_{i_4}(t+\tau_3) dt - \frac{1}{2T} \int_{-T}^{T} g_{i_1}(t) g_{i_2}(t+\tau_1) \right.$$

$$\times dt \frac{1}{2T} \int_{-T}^{T} g_{i_3}(t+\tau_2) g_{i_4}(t+\tau_3) dt - \frac{1}{2T} \int_{-T}^{T} g_{i_1}(t) g_{i_3}(t+\tau_2) dt \frac{1}{2T} \int_{-T}^{T} g_{i_2}(t+\tau_1) g_{i_4}(t+\tau_3) dt$$

$$\left. - \frac{1}{2T} \int_{-T}^{T} g_{i_1}(t) g_{i_4}(t+\tau_3) dt \frac{1}{2T} \int_{-T}^{T} g_{i_2}(t+\tau_1) g_{i_3}(t+\tau_2) dt \right\}, \tag{2}$$

and the fourth-order cumulant spectrum

$$C_4(\omega_1, \omega_2, \omega_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c_4(\tau_1, \tau_2, \tau_3) e^{-j(\omega_1\tau_1 + \omega_2\tau_2 + \omega_3\tau_3)} d\tau_1 \, d\tau_2 \, d\tau_3 \tag{3}$$

reduces to

$$c_4(\omega_1, \omega_2, \omega_3) = \sum_{i_1=1}^{N_t} \sum_{i_2=1}^{N_t} \sum_{i_3=1}^{N_t} \sum_{i_4=1}^{N_t} \left\{ \frac{1}{2T} G_{i_1}(-(\omega_1+\omega_2+\omega_3)) G_{i_2}(\omega_1) G_{i_3}(\omega_2) G_{i_4}(\omega_3) - \frac{1}{2T} G_{i_1}(-\omega_1) \right.$$

$$\times G_{i_2}(\omega_1) G_{i_3}(\omega_2) G_{i_4}(\omega_3) \frac{\sin((\omega_2+\omega_3)T)}{(\omega_2+\omega_3)T} - \frac{1}{2T} G_{i_1}(-\omega_2) G_{i_3}(\omega_2) G_{i_2}(\omega_1) G_{i_4}(\omega_3) \frac{\sin((\omega_1+\omega_3)T)}{(\omega_1+\omega_3)T}$$

$$\left. - \frac{1}{2T} G_{i_1}(-\omega_3) G_{i_4}(\omega_3) G_{i_2}(\omega_1) G_{i_3}(\omega_2) \frac{\sin((\omega_1+\omega_2)T)}{(\omega_1+\omega_2)T} \right\}, \tag{4}$$

where, assuming T was sufficiently large, the following approximation was used above:

$$G(\omega) \approx \int_{-T}^{T} g(t) e^{-j\omega t} \, dt. \tag{5}$$

Since, also, for T sufficiently large,

$$\frac{\sin((\omega_\alpha + \omega_\beta)T)}{(\omega_\alpha + \omega_\beta)T} = \begin{cases} 1, & \text{if } \omega_\alpha = -\omega_\beta, \\ 0, & \text{otherwise}, \end{cases} \tag{6}$$

and, for real data, the result simplifies to,

$$c_4(\omega_1, \omega_2, \omega_3) = \sum_{i_1=1}^{N_t} \sum_{i_2=1}^{N_t} \sum_{i_3=1}^{N_t} \sum_{i_4=1}^{N_t} \left\{ \frac{1}{2T} G_{i_1}^*((\omega_1+\omega_2+\omega_3)) G_{i_2}(\omega_1) G_{i_3}(\omega_2) G_{i_4}(\omega_3) \right.$$

$$- \frac{1}{2T} G_{i_1}^*(\omega_1) G_{i_2}(\omega_1) G_{i_3}^*(\omega_3) G_{i_4}(\omega_3) - \frac{1}{2T} G_{i_1}^*(\omega_2) G_{i_3}(\omega_2) G_{i_2}^*(\omega_3) G_{i_4}(\omega_3)$$

$$\left. - \frac{1}{2T} G_{i_1}^*(\omega_3) G_{i_4}(\omega_3) G_{i_2}^*(\omega_1) G_{i_3}(\omega_1) \right\}. \tag{7}$$

This is the form of the fourth-order cumulant spectrum that shall be evaluated, except for the scale factor $(1/2T)$, at the output of a beamformer and compared with the output spectrum in the next section. In the above equation, the first part on the right is called the fourth-order moment spectrum, denoted by $M_4(\omega_1, \omega_2, \omega_3)$, and the other three terms are products of second-order terms, $G_i(\omega_\alpha) G_{ii}^*(\omega_\alpha)$, $i, ii = 1, ..., N_t$, $\alpha = 1, 2, 3$. Given the fourth-order cumulant spectrum, the fourth-order moment spectrum can be recovered as well as the other terms if $G_i(\omega_c) = 0$, $i = 1, ..., N_t$, for some ω_c, as

$$M_4(\omega_1, \omega_2, \omega_3) = \sum_{i_1=1}^{N_t} \sum_{i_2=1}^{N_t} \sum_{i_3=1}^{N_t} \sum_{i_4=1}^{N_t} \frac{1}{2T} G_{i_1}^*((\omega_1+\omega_2+\omega_3)) G_{i_2}(\omega_1) G_{i_3}(\omega_2) G_{i_4}(\omega_4) = C_4(\omega_1, \omega_2, \omega_3)$$

$$- C_4(\omega_1, \omega_2, \omega_c) - C_4(\omega_1, \omega_c, \omega_3) - C_4(\omega_c, \omega_2, \omega_3). \tag{8}$$

The application determines whether the fourth-order moment spectrum $M_4(\omega_1, \omega_2, \omega_3)$ or the fourth-order cumulant spec-
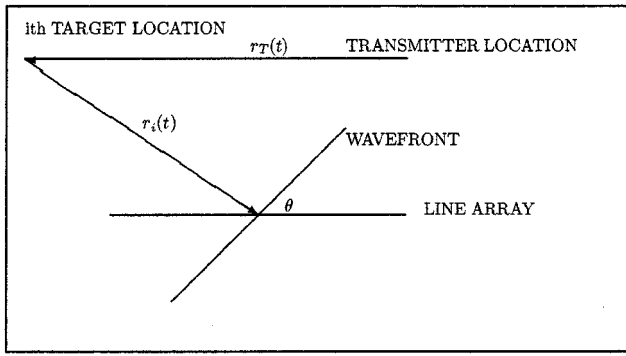
FIG. 1. Transmitter, target, and line array geometry.

trum $C_4(\omega_1,\omega_2,\omega_3)$ is used. Even though the phase of $F(\omega)$ cannot be recovered from $|F(\omega)|^2$, the phase of $F(\omega)$ can be recovered from the fourth-order moment spectrum. This is based on the work of Brillenger who developed a recursive algorithm to recover the phase of $F(\omega)$ from the bispectrum[9] and Matsuoka and Ulrych[10] extended Brillenger's algorithm to recover the phase from the trispectrum.[11–14] In the next section the implications of recovering phase from beamformed data is discussed. This is mentioned here because phase information can be important in classification applications.[15–19]

I. BEAMFORMED DATA

Let a target be insonified by an active sonar pulsed waveform from a transmitter and let its reflected plane wave return intersect a line array as shown in Fig. 1. Of course, the transmitter, target, and line array need not be in the same plane in general. The objectives of an active sonar system are to determine the arrival angle with respect to the array orientation, the range of the target with respect to a reference point on the array, and the target characteristics, from the returned signal. Often, active sonar systems match the returned signal with the transmitted waveform. This approach will be compared with the fourth-order cumulant spectrum processing approach for wide bandwidth transmitted waveforms.[15]

The output of a ''delay and sum'' beamformer from the one-dimensional line array with uniformly distributed sensors is

$$y(t)=\sum_{q=0}^{N_b-1}\sum_{i=1}^{N_t} f_i\left(t-\frac{2r_i(t)}{c}-qd(\sin(\theta_i)-\sin(\theta_s))/c\right),$$
(9)

where N_b is the number of sensors in the array, d is the sensor separation, c is the speed of sound, $r_i(t)$ is the instantaneous range to the ith target, N_t is the number of targets insonified, θ_i is the signal arrival direction of the ith target, and θ_s is the direction that the array is steered to. Amplitude scaling from propagation loss has been included in the returned signal, f_i. Since, here, the range to any target is much longer then the distance separating the transmitter and array it will be assumed that they are nearly colocated and therefore $r_T(t)\approx r_i(t)$. It was also assumed that the signal is

propagating in the plane of the array and array steering is confined to the plane of the array. Out of plane signal arrivals would be handled in a similar way as here but will not be explicitly indicated in the equations. Other array configurations, i.e., two-dimensional planar arrays,[21] are easily incorporated in the results but are not discussed further. In the following the instantaneous range will be approximated as

$$r_i(t)\approx r_i(0)+t\dot{r}_i(0),$$
(10)

where, $r_i(0)$ is the initial range to the ith target and $\dot{r}_i(0)$ is the constant velocity of the ith target. Inserting this approximation into $y(t)$, its Fourier transform is given by

$$Y(\omega,\Theta,\theta_s)=\int_{-\infty}^{\infty} y(t)e^{-j\omega t}\,dt=\sum_{i=1}^{N_t} G_i(\omega),$$
(11)

where

$$G_i(\omega)=(1-M_{r_i})F_i(\omega(1-M_{r_i}))$$
$$\times e^{-j\omega(1-M_{r_i})A(\theta_i,\theta_s)(N_b-1)d/2c}$$
$$\times e^{-j\omega(1-M_{r_i})[2r_i(0)/c]}$$
$$\times \frac{\sin(\omega(1-M_{r_i})A(\theta_i,\theta_s)N_bd/2c)}{\sin(\omega(1-M_{r_i})A(\theta_i,\theta_s)d/2c)},$$
(12)

$A(\theta_i,\theta_s)=\sin(\theta_i)-\sin(\theta_s)$, Θ is the set of target arrival angles, and $1-M_{r_i}=1/(1-2\dot{r}_i(0)/c)$ represents the effect of the ith target motion on the received signal.[13,20] It will be convenient to define the following relationship:

$$D(\omega A(\theta_i,\theta_s))=\frac{\sin(\omega(1-M_{r_i})A(\theta_i,\theta_s)N_bd/2c)}{\sin(\omega(1-M_{r_i})A(\theta_i,\theta_s)d/2c)}.$$
(13)

The phase of $Y(\omega,\Theta,\theta_s)$ is formally defined as

$$\phi_{Y(\omega,\Theta,\theta_s)}=\tan^{-1}\left(\frac{\text{Im}[Y(\omega,\Theta,\theta_s)]}{\text{Re}[Y(\omega,\Theta,\theta_s)]}\right),$$
(14)

where Im is the imaginary part and Re is the real part. Whereas, the phase of the ith target at the beamformer output is

$$\phi_{F_i}(\omega(1-M_{r_i}))-\omega(1-M_{r_i})A(\theta_i,\theta_s)(N_b-1)d/2c$$
$$-\omega(1-M_{r_i})\frac{2r_i(0)}{c}+\phi_a(\omega),$$
(15)

where $\phi_{F_i}(\omega(1-M_{r_i}))$, is the phase of $F_i(\omega(1-M_{r_i}))$, and $\phi_a(\omega)=\pi$, if $D(\omega A(\theta_i,\theta_s=\theta_i))<0$ and zero otherwise. From $Y(\omega,\Theta,\theta_s)$, therefore, the phase of $F_i(\omega(1-M_{r_i}))$ can only be recovered exactly if $\theta_s=\theta_i$ and the contributions from the other targets are negligible after beamforming. Under the condition $\theta_s=\theta_i$, $Y(\omega,\Theta,\theta_s=\theta_i)=N_b(1-M_{r_i})F_i(\omega(1-M_{r_i}))e^{-j\omega(1-M_{r_i})[2r_i(0)/c]}$, assuming contributions from the other targets after beamforming are negli-

gible. As θ_s is steered over the possible arrival directions grating lobes and beam broadening will occur if the beamformer is used over wide bandwidths.[21] In some beamformer designs the steering directions are preselected and the measured phase of $F_i(\omega(1-M_{r_i}))$ at these steering directions may well be distorted, as can be seen from Eq. (15), even if the contributions of the other targets are negligible. As will be shown below the phase of $F_i(\omega(1-M_{r_i}))$ can be recovered exactly at the output of the beamformer for steering directions in the vicinity of the ith target from the fourth-order moment spectrum except for a phase function that is either 0 or π.

These results can be stated in a theorem and its corollary:

Theorem 1:

$$\lim_{N_b \to \infty} Y(\omega,\Theta,\theta_s \not\subset \Theta)/N_b \not\to 0, \quad \forall \omega, \quad -\frac{\pi}{2} \leq \theta_s \leq \frac{\pi}{2}. \tag{16}$$

Corollary:

$$\lim_{N_b \to \infty} Y(\omega,\Theta,\theta_s = \theta_k)/N_b$$
$$\to (1-M_{r_k})F_k(\omega(1-M_{r_k}))$$
$$\times e^{-\omega(1-M_{r_k})[2r_k(0)/c]}, \quad \forall \omega, \tag{17}$$

for $i \neq k$, $A(\theta_i,\theta_k) \neq 2\pi I/K(1-M_{r_i})d$, and $I = \pm 1, \pm 2,...$. Theorem 1 states that the output of the normalized beamformer does not approach zero for all steering directions away from the targets for all ω even as the number of sensors become large. However, for a steering direction pointing towards the kth target, $Y(\omega,\Theta,\theta_s = \theta_k)/N_b$ approaches the kth target's Fourier transform at the normalized beamformer output if no grating lobes occur from the other target directions (i.e., $i \neq k$).

Proof of Theorem 1: The normalized beamformer output is

$$Y(\omega,\Theta,\theta_s \not\subset \Theta)/N_b = \sum_{i=1}^{N_t} G_i(\omega)/N_b. \tag{18}$$

For Theorem 1 to be true there must be multiple solutions to the beampattern [i.e., $BP_{so}(A) = 1$], where

$$BP_{so}(A) = \left| \frac{\sin[K(1-M_{r_i})A(\theta_i,\theta_s)N_b d/2]}{N_b \sin[K(1-M_{r_i})A(\theta_i,\theta_s)d/2]} \right|, \tag{19}$$

and $K = \omega/c$ is the acoustic wave number.[20]

The solutions for $BP_{so} = 1$ are $\sin[K(1-M_{r_i})A(\theta_i,\theta_s)d/2] = 0$ or when $A(\theta_i,\theta_s) = 2\pi I/K(1-M_{r_i})d$, $I = 0,\pm 1,\pm 2,...$. To prevent grating lobes or multiple solutions from occurring, (i.e., $I = \pm 1, \pm 2,...$), over all steering directions the element spacing d must be $d < \lambda/2$, where λ is the wavelength of the signal being beamformed.[21] When the array is receiving signals below its

design wavelength, λ_d, the element spacing will be $d = \lambda_d/2 > \lambda/2$, and grating lobes will occur within the "visible" region of the beamformer. On the other hand for low wave numbers, such that $K(1-M_{r_i})A(\theta,\theta_s)N_b d/2 < 1$, $BP_{so}(A) \approx 1$ for all steering directions. Both of these effects can be controlled with the fourth-order spectrum. The corollary to Theorem 1 implies that when the array it steered toward the kth target and there are no multiple solutions caused from the other target locations, (i.e., $i \neq k$), then the beamformer output approaches the Fourier transform of the kth target. This completes the proof for Theorem 1 and its corollary.

The fourth-order cumulant spectrum at the output of the beamformer follows as

$$C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$$
$$= Y(\omega_1,\Theta,\theta_s)Y(\omega_2,\Theta,\theta_s)Y(\omega_3,\Theta,\theta_s)Y^*((\omega_1+\omega_2$$
$$+\omega_3),\Theta,\theta_s) - |Y(\omega_1,\Theta,\theta_s)|^2|Y(\omega_2,\Theta,\theta_s)|^2$$
$$- |Y(\omega_1,\Theta,\theta_s)|^2|Y(\omega_3,\Theta,\theta_s)|^2$$
$$- |Y(\omega_2,\Theta,\theta_s)|^2|Y(\omega_3,\Theta,\theta_s)|^2, \tag{20}$$

which reduces to

$$C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$$
$$= \sum_{i_1=1}^{N_t}\sum_{i_2=1}^{N_t}\sum_{i_3=1}^{N_t}\sum_{i_4=1}^{N_t} G_{i_1}(\omega_1)G_{i_2}(\omega_2)G_{i_3}(\omega_3)G_{i_4}^*(\omega_4)$$
$$- \sum_{i_1=1}^{N_t}\sum_{i_2=1}^{N_t}\sum_{i_3=1}^{N_t}\sum_{i_4=1}^{N_t} G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)G_{i_3}(\omega_2)G_{i_4}^*(\omega_2)$$
$$- \sum_{i_1=1}^{N_t}\sum_{i_2=1}^{N_t}\sum_{i_3=1}^{N_t}\sum_{i_4=1}^{N_t} G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)G_{i_3}(\omega_3)G_{i_4}^*(\omega_3)$$
$$- \sum_{i_1=1}^{N_t}\sum_{i_2=1}^{N_t}\sum_{i_3=1}^{N_t}\sum_{i_4=1}^{N_t} G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)G_{i_3}(\omega_3)G_{i_4}^*(\omega_3), \tag{21}$$

where $\omega_4 = \omega_1 + \omega_2 + \omega_3$. The Appendix gives the general expression for N_t targets. Here the result is given for two targets, i.e., $N_t = 2$. The fourth-order cumulant spectrum at the output of the beamformer for two targets follows as

$$C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$$

$$= (G_1(\omega_1)G_1(\omega_2)G_1(\omega_3)G_1^*(\omega_4) - |G_1(\omega_1)|^2|G_1(\omega_2)|^2 - |G_1(\omega_1)|^2|G_1(\omega_3)|^2 - |G_1(\omega_2)|^2|G_1(\omega_3)|^2)$$

$$+ (G_2(\omega_1)G_2(\omega_2)G_2(\omega_3)G_2^*(\omega_4) - |G_2(\omega_1)|^2|G_2(\omega_2)|^2 - |G_2(\omega_1)|^2|G_2(\omega_3)|^2 - |G_2(\omega_2)|^2|G_2(\omega_3)|^2)$$

$$+ (G_2(\omega_1)G_1(\omega_2)G_1(\omega_3)G_1^*(\omega_4) - G_2(\omega_1)G_1^*(\omega_1)|G_1(\omega_2)|^2 - G_2(\omega_1)G_1^*(\omega_1)|G_1(\omega_3)|^2$$

$$- G_2(\omega_2)G_1^*(\omega_2)|G_1(\omega_3)|^2)$$

$$+ (G_1(\omega_1)G_2(\omega_2)G_1(\omega_3)G_1^*(\omega_4) - G_1(\omega_1)G_2^*(\omega_1)|G_1(\omega_2)|^2 - G_1(\omega_1)G_2^*(\omega_1)|G_1(\omega_3)|^2$$

$$- G_1(\omega_2)G_2^*(\omega_2)|G_1(\omega_3)|^2)$$

$$+ (G_2(\omega_1)G_2(\omega_2)G_1(\omega_3)G_1^*(\omega_4) - |G_2(\omega_1)|^2|G_1(\omega_2)|^2 - |G_2(\omega_1)|^2|G_1(\omega_3)|^2 - |G_2(\omega_2)|^2|G_1(\omega_3)|^2)$$

$$+ (G_1(\omega_1)G_1(\omega_2)G_2(\omega_3)G_1^*(\omega_4) - |G_1(\omega_1)|^2G_2(\omega_2)G_1^*(\omega_2) - |G_1(\omega_1)|^2G_2(\omega_3)G_1^*(\omega_3)$$

$$- |G_1(\omega_2)|^2G_2(\omega_3)G_1^*(\omega_3))$$

$$+ (G_2(\omega_1)G_1(\omega_2)G_2(\omega_3)G_1^*(\omega_4) - G_2(\omega_1)G_1^*(\omega_1)G_2(\omega_2)G_1^*(\omega_2) - G_2(\omega_1)G_1^*(\omega_1)G_2(\omega_3)G_1^*(\omega_3)$$

$$- G_2(\omega_2)G_1^*(\omega_2)G_2(\omega_3)G_1^*(\omega_3))$$

$$+ (G_1(\omega_1)G_2(\omega_2)G_1(\omega_3)G_2^*(\omega_4) - G_1(\omega_1)G_2^*(\omega_1)G_2(\omega_2)G_1^*(\omega_2) - G_1(\omega_1)G_2^*(\omega_1)G_2(\omega_3)G_1^*(\omega_3)$$

$$- G_1(\omega_2)G_2^*(\omega_2)G_2(\omega_3)G_1^*(\omega_3))$$

$$+ (G_2(\omega_1)G_2(\omega_2)G_2(\omega_3)G_1^*(\omega_4) - |G_2(\omega_1)|^2G_2(\omega_2)G_1^*(\omega_2) - |G_2(\omega_1)|^2G_2(\omega_3)G_1^*(\omega_3)$$

$$- |G_2(\omega_2)|^2G_2(\omega_3)|G_1^*(\omega_3))$$

$$+ (G_1(\omega_1)G_1(\omega_2)G_1(\omega_3)G_2^*(\omega_4) - |G_1(\omega_1)|^2G_2(\omega_2)G_1^*(\omega_2) - |G_1(\omega_1)|^2G_2(\omega_3)G_1^*(\omega_3)$$

$$- |G_1(\omega_2)|^2G_2(\omega_3)G_1^*(\omega_3))$$

$$+ (G_2(\omega_1)G_1(\omega_2)G_1(\omega_3)G_2^*(\omega_4) - G_2(\omega_1)G_1^*(\omega_1)G_2(\omega_2)G_1^*(\omega_2) - G_2(\omega_1)G_1^*(\omega_1)G_2(\omega_3)G_1^*(\omega_3)$$

$$- G_2(\omega_2)G_1^*(\omega_2)G_2(\omega_3)G_1^*(\omega_3))$$

$$+ (G_1(\omega_1)G_2(\omega_2)G_1(\omega_3)G_2^*(\omega_4) - G_1(\omega_1)G_2^*(\omega_1)G_2(\omega_2)G_1^*(\omega_2) - G_1(\omega_1)G_2^*(\omega_1)G_1^*(\omega_3)G_2(\omega_3)$$

$$- G_1(\omega_2)G_2^*(\omega_2)G_1^*(\omega_3)G_2(\omega_3))$$

$$+ (G_2(\omega_1)G_2(\omega_2)G_1(\omega_3)G_2^*(\omega_4) - |G_2(\omega_1)|^2G_2(\omega_2)G_1^*(\omega_2) - |G_2(\omega_1)|^2G_2(\omega_3)G_1^*(\omega_3)$$

$$- |G_2(\omega_2)|^2G_2(\omega_3)G_1^*(\omega_3))$$

$$+ (G_1(\omega_1)G_1(\omega_2)G_2(\omega_3)G_2^*(\omega_4) - |G_1(\omega_1)|^2|G_2(\omega_2)|^2 - |G_1(\omega_1)|^2|G_2(\omega_3)|^2 - |G_1(\omega_2)|^2|G_2(\omega_3)|^2)$$

$$+ (G_2(\omega_1)G_1(\omega_2)G_2(\omega_3)G_2^*(\omega_4) - G_2(\omega_1)G_1^*(\omega_1)|G_2(\omega_2)|^2 - G_2(\omega_1)G_1^*(\omega_1)|G_2(\omega_3)|^2$$

$$- G_2(\omega_2)G_1^*(\omega_2)|G_2(\omega_3)|^2)$$

$$+ (G_1(\omega_1)G_2(\omega_2)G_2(\omega_3)G_2^*(\omega_4) - G_1(\omega_1)G_2^*(\omega_1)|G_2(\omega_2)|^2 - G_1(\omega_1)G_2^*(\omega_1)|G_2(\omega_3)|^2$$

$$- G_1(\omega_2)G_2^*(\omega_2)|G_2(\omega_3)|^2). \tag{22}$$

For the terms on the right of the above equation, the first term is the fourth-order cumulant spectrum for the first target and the second term is the fourth-order cumulant spectrum for the second target. These terms are called target terms. The other terms above are mixtures from both targets and are called mixture terms. For the general result, for N_t targets as shown in the Appendix, the first N_t terms are the target terms and the rest are mixture terms comprising combinations from all of the targets. When the array is steered to a particular target, say the kth, the other terms in the above expression and in the general expression given in the Appendix will be suppressed by the beamformer depending on the mixture. Maximum suppression occurs when a term is composed of elements from the same target, say the ith, but the array is steered to the kth target, i.e., $\theta_s = \theta_k$. The terms $e^{-j\omega(1-M_{r_i})A(\Theta,\theta_s)(N_b+1)/2}$ and $e^{-j\omega(1-M_{r_i})[2r_i(0)/c]}$ occurring

in $Y(\omega,\Theta,\theta_s)$ will be cancelled in $C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$ for all target terms due to the structure of the fourth-order cumulant spectrum. This has an advantage in extracting the phase of $F_k(\omega(1-M_{r_k}))$ as will be discussed below but this

also eliminates the possibility of extracting range directly from $C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$, as can be done from $Y(\omega,\Theta,\theta_s)$. However, range can be deduced indirectly.[22] When $\theta_s=\theta_k$,

$$C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s=\theta_k)\approx N_b^4(1-M_{r_k})^4 F_k(\omega_1(1-M_{r_k}))F_k(\omega_2(1-M_{r_k}))F_k(\omega_3(1-M_{r_k}))F_k^*[(\omega_1+\omega_2+\omega_3)$$
$$\times(1-M_{r_k})]-N_b^4(1-M_{r_k})^4|F_k(\omega_1(1-M_{r_k}))|^2|F_k(\omega_2(1-M_{r_k}))|^2-N_b^4(1-M_{r_k})^4|F_k(\omega_1(1-M_{r_k}))|^2|F_k(\omega_3(1-M_{r_k}))|^2-N_b^4(1-M_{r_k})^4|F_k(\omega_2(1-M_{r_k}))|^2|F_k(\omega_3(1-M_{r_k}))|^2, \qquad (23)$$

where the other target terms, i.e., $\theta_i\neq\theta_k$, and the mixture terms were assumed sufficiently suppressed by the beamformer.

The fourth-order moment spectrum, $M_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$, can also be recovered from $C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$ if $F_i(\omega_c)=0$, $i=1,...,N_t$, for some ω_c. The phase from $M_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s\to\theta_k)$ is

$$\psi(\omega_1,\omega_2,\omega_3,\theta_k)=(\phi_{F_k}(\omega_1(1-M_{r_k}))+\phi_{F_k}(\omega_2(1-M_{r_k}))+\phi_{F_k}(\omega_3(1-M_{r_k}))$$
$$-\phi_{F_k}((\omega_1+\omega_2+\omega_3)(1-M_{r_k}))$$
$$+\phi_4(\omega_1,\omega_2,\omega_3)), \qquad (24)$$

where $\phi_4(\omega_1,\omega_2,\omega_3)=\pi$ if

$$(D(\omega_1 N_b)D(\omega_2 N_b)D(\omega_3 N_b)D((\omega_1+\omega_2+\omega_3)N_b))<0$$

and zero otherwise. Therefore the phase $\phi_{F_k}(\omega(1-M_{r_k}))$ can be recovered using Brillenger's modified recursive algorithm[8,9] for steering directions in the vicinity of the kth target (i.e., $\theta_s\to\theta_k$), except for jumps in phase equal to π from $\phi_4(\omega_1,\omega_2,\omega_3)$. These jumps in π can be detected and corrected in a similar way as discussed in Ref. 23. When $\theta_s=\theta_k$, then $\phi_4(\omega_1,\omega_2,\omega_3)=0$.

The above discussion for the fourth-order cumulant spectrum at the beamformer output can be stated in a theorem and corollary.

Theorem 2:

$$\lim_{N_b\to\infty} C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s\not\subset\Theta)/N_b^4$$
$$\to 0, \quad \forall(\omega_1,\omega_2,\omega_3), \quad -\pi/2\leq\theta_s\leq\pi/2, \qquad (25)$$
if $A(\theta_i,\theta_s\not\subset\Theta)\neq 2\pi I/\alpha(1-M_{r_i})d$, $I=\pm 1,\pm 2,...$, i

$=1,...,N_t$, for at least one $\alpha=K_1,K_2,K_3$, or when, $\alpha=K_1+K_2+K_3$, where K_i, $i=1,2,3$ are the wave numbers associated with the fourth-order spectrum.

Corollary:

$$\lim_{N_b\to\infty} C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s=\theta_k)/N_b^4$$
$$\to (1-M_{r_k})^4 F_k(\omega_1(1-M_{r_k}))F_k(\omega_2(1-M_{r_k}))$$
$$\times F_k(\omega_3(1-M_{r_k}))F_k^*[(\omega_1+\omega_2+\omega_3)(1-M_{r_k})],$$
$$-(1-M_{r_k})^4|F_k(\omega_1(1-M_{r_k}))|^2|F_k(\omega_2(1-M_{r_k}))|^2,$$
$$-(1-M_{r_k})^4|F_k(\omega_1(1-M_{r_k}))|^2|F_k(\omega_3(1-M_{r_k}))|^2,$$
$$-(1-M_{r_k})^4|F_k(\omega_2(1-M_{r_k}))|^2|F_k(\omega_3(1-M_{r_k}))|^2, \qquad (26)$$

if $A(\theta_i,\theta_s=\theta_k)\neq 2\pi I/\alpha(1-M_{r_i})d$, $I=\pm 1,\pm 2,...$, $i\neq k$, for at least one $\alpha=K_1,K_2,K_3$ or when $\alpha=K_1+K_2+K_3$.

Theorem 2 states that the normalized fourth-order cumulant spectrum at the beamformer output approaches zero for steering directions away from the targets as the number of sensors in the array becomes large. Whereas, the corollary states that for a steering direction pointing toward a target the normalized fourth-order cumulant spectrum at the beamformer output approaches the fourth-order cumulant spectrum of the target pointed towards.

Proof of Theorem 2: The reasons why the fourth-order cumulant spectrum reduces the deleterious effects of wide bandwidths is easily shown by examining its spatial properties. The fourth-order moment spectrum's spatial beampattern is defined as follows:

$BP_{fo}(A)$

$$=\left|\frac{\sin[K_1(1-M_{r_{i_1}})A(\theta_{i_1},\theta_s)N_b d/2]}{N_b\sin[K_1(1-M_{r_{i_1}})A(\theta_{i_1},\theta_s)d/2]}\frac{\sin[K_2(1-M_{r_{i_2}})A(\theta_{i_2},\theta_s)N_b d/2]}{N_b\sin[K_2(1-M_{r_{i_2}})A(\theta_{i_2},\theta_s)d/2]}\frac{\sin[K_3(1-M_{r_{i_3}})A(\theta_{i_3},\theta_s)N_b d/2]}{N_b\sin[K_3(1-M_{r_{i_3}})A(\theta_{i_3},\theta_s)d/2]}\frac{\sin[(K_1+K_2+K_3)(1-M_{r_{i_4}})A(\theta_{i_4},\theta_s)N_b d/2]}{N_b\sin[(K_1+K_2+K_3)(1-M_{r_{i_4}})A(\theta_{i_4},\theta_s)d/2]}\right|.$$

$$(27)$$

The other term's beampatterns are

$$BP_{go}(A)$$

$$= \left| \frac{\sin[K_p(1-M_{r_{i_1}})A(\theta_{i_1},\theta_s)N_bd/2]}{N_b\sin[K_p(1-M_{r_{i_1}})A(\theta_{i_1},\theta_s)d/2]} \frac{\sin[K_p(1-M_{r_{i_2}})A(\theta_{i_2},\theta_s)N_bd/2]}{N_b\sin[K_p(1-M_{r_{i_2}})A(\theta_{i_2},\theta_s)d/2]} \frac{\sin[K_q(1-M_{r_{i_3}})A(\theta_{i_3},\theta_s)N_bd/2]}{N_b\sin[K_q(1-M_{r_{i_3}})A(\theta_{i_3},\theta_s)d/2]} \frac{\sin[K_q(1-M_{r_{i_4}})A(\theta_{i_4},\theta_s)N_bd/2]}{N_b\sin[K_q(1-M_{r_{i_4}})A(\theta_{i_4},\theta_s)d/2]} \right|,$$

(28)

where $p,q=1,2,3$, $p \neq q$.

In order for $BP_{fo}(A)=1$, the following solutions must occur simultaneously: $\{\sin[K_1(1-M_{r_{i_1}})A(\theta_{i_1},\theta_s)d/2]=0$, $\sin[K_2(1-M_{r_{i_2}})A(\theta_{i_2},\theta_s)d/2]=0$, $\sin[K_3(1-M_{r_{i_3}})A(\theta_{i_3},\theta_s)d/2]=0$, and $\sin[(K_1+K_2+K_3)(1-M_{r_{i_4}})A(\theta_{i_4},\theta_s)d/2]=0\}$, or equivalently, the following must occur simultaneously: $\{A(\theta_{i_1},\theta_s)=2\pi I_1/K_1(1-M_{r_{i_1}})d$, $I_1=0,\pm1,\pm2,\ldots$, $A(\theta_{i_2},\theta_s)=2\pi I_2/K_2(1-M_{r_{i_2}})d$, $I_2=0,\pm1,\pm2,\ldots$, $A(\theta_{i_3},\theta_s)=2\pi I_3/K_3(1-M_{r_{i_3}})d$, $I_3=0,\pm1,\pm2,\ldots$, $A(\theta_{i_4},\theta_s)=2\pi I_4/(K_1+K_2+K_3)(1-M_{r_{i_4}})d$, $I_4=0,\pm1,\pm2,\ldots\}$, $i_1,i_2,i_3,i_4=1,\ldots,N_t$. So that, if any, $A(\theta_i,\theta_s\not\subseteq\Theta)\neq 2\pi I/\alpha(1-M_{r_i})d$, $I=\pm1,\pm2,\ldots$, $i=1,\ldots,N_t$, for at least one $\alpha=K_1,K_2,K_3$, or when $\alpha=K_1+K_2+K_3$ solutions $BP_{fo}(A=0)\neq1$, cannot occur. In addition, for low wave numbers, $BP_{fo}(A)=1$, only for $A=0$.

For $BP_{go}(A)=1$, the following must occur simultaneously, $\{\sin[K_p(1-M_{r_{i_1}})A(\theta_{i_1},\theta_s)d/2]=0, \sin[K_p(1-M_{r_{i_2}})\times A(\theta_{i_2},\theta_s)d/2]=0$, $\sin[K_q(1-M_{r_{i_3}})A(\theta_{i_3},\theta_s)d/2]=0$, and $\sin[(K_q)(1-M_{r_{i_4}})A(\theta_{i_4},\theta_s)d/2]=0\}$, or equivalently, the following must occur simultaneously, $\{A(\theta_{i_1},\theta_s)=2\pi I_1/K_p(1-M_{r_{i_1}})d$, $I_1=0,\pm1,\pm2,\ldots$, $A(\theta_{i_2},\theta_s)=2\pi I_2/K_p(1-M_{r_{i_2}})d$, $I_2=0,\pm1,\pm2,\ldots$, $A(\theta_{i_3},\theta_s)=2\pi I_3/K_q(1-M_{r_{i_3}})d$, $I_3=0,\pm1,\pm2,\ldots$, $A(\theta_{i_4},\theta_s)=2\pi I_4/(K_q)(1-M_{r_{i_4}})d$, $I_4=0,\pm1,\pm2,\ldots\}$ for $p,q=1,2,3$, $p\neq q$, $i_1,i_2,i_3,i_4=1,\ldots,N_t$. So if any, $A(\theta_i,\theta_s=\theta_k)\neq 2\pi I/\alpha(1-M_{r_i})d$, $I=\pm1,\pm2,\ldots$, $i\neq k$, for at least one $\alpha=K_1,K_2,K_3$ or when $\alpha=K_1+K_2+K_3$ solutions $BP_{go}(A\neq0)=1$, cannot occur. Therefore there are no solutions other then $A=0$. It should be pointed out that Theorem 2 only guarantees that $BP_{fo}\neq1$ and $BP_{go}\neq1$ for $A\neq0$, and therefore does not specify the actual values of BP_{fo} and BP_{go}. The corollary follows since $A(\theta_i,\theta_s=\theta_k)\neq 2\pi I/\alpha(1-M_{r_i})d$, $I=\pm1,\pm2,\ldots$, $i=k$, for at least one $\alpha=K_1,K_2,K_3$ or when $\alpha=K_1+K_2+K_3$, therefore all mixtures terms and all other target terms (i.e., $i\neq k$) will converge to zero as $N_b\to0$. This completes the proof for Theorem 2 and its corollary.

A comparision of the matched filtered beamformer output $Y(\omega(1-M_r),\Theta,\theta_s)X^*(\omega)$ and the fourth-order cumulant spectrum output $C(\omega(1-M_r),\omega_a,\omega_b,\Theta,\theta_s)$, where $\omega(1-M_r)=\omega_1(1-M_r)$ and the two other frequencies ω_a and ω_b are preselected and fixed throughout, will be given below. The matched filtered signal $X(\omega)$ represents the linear frequency modulated (LFM)[24,25] transmitted waveform for zero Doppler, i.e., $M_r=0$. The actual values of ω_a and ω_b are important and effect the results. They were chosen

under the constraint $\omega_a+\omega_b=2\pi B$, where B is the bandwidth and then fixed for all comparisons. The simulated results will include two targets at different arrival directions but at the same velocity. Two cases will be considered. The first case is for stationary, zero Doppler targets with respect to the arrival direction. This case matches the transmitted waveform and is therefore the matched filter. The second case is for 28-kn targets with respect to the arrival direction. This mismatches the transmitted signal and the matched filter. Other transmitted waveforms may have better mismatch performance than the LFM waveform,[26] but the LFM's spectral shape is important for wide bandwidths.

The array is linear with uniform $\lambda/2$ sensor spacing and a fixed λ. For the results to follow the linear array has ten elements uniformly separated consistent with a frequency of 4160 Hz, which is about 0.18 m for a sound speed of 1500 m/s. There will be two sets of data. The first set of data is from a perfectly reflecting target so the returned signal is the transmitted waveform modified by the target motion with respect to the receiver. The arrival directions are 0° and 45°. Two velocities, 0 and 28 kn with respect to the receiver are compared. The first velocity represents the matched case and the second velocity the mismatched case. The second set of data is from five elastic spheres:[27,28] rigid, Lucite, steel, aluminum, and nickel. Sheres are an important underwater acoustic target and have even been used as a calibration to compare with the sonar of dolphins.[29] Each of the sphere's impulse response is convolved with the transmitted waveform and propagated toward the array. These targets will also have two velocities of 0 and 28 kn and arrival directions of 0° and 45°. Therefore $F(\omega(1-M_r))=H(\omega(1-M_r))X(\omega(1-M_r))$, where $H(\omega(1-M_r))$ is the motion modified transfer function of the sphere[15] and $X(\omega(1-M_r))$ is the Fourier transform of the motion modified transmitted waveform. This result can be shown by defining the transmitted signal convolved with the target, at the target location, as

$$f\left(t-\frac{r_T(t)}{c}\right)=\int_{-\infty}^{\infty}h(\tau)x\left(t-\frac{r_T(t)}{c}-\tau\right)d\tau,$$

(29)

where the geometry is defined in Fig. 1. At the array location, assuming $r_T(t)=r(t)$,

$$f\left(t-\frac{r_T(t)}{c}-\frac{r(t)}{c}\right)=f\left(t-\frac{2r(t)}{c}\right)$$

$$=\int_{-\infty}^{\infty}h(\tau)x\left(t-\frac{2r(t)}{c}-\tau\right)d\tau.$$

(30)

Taking the Fourier transform of both sides of the above
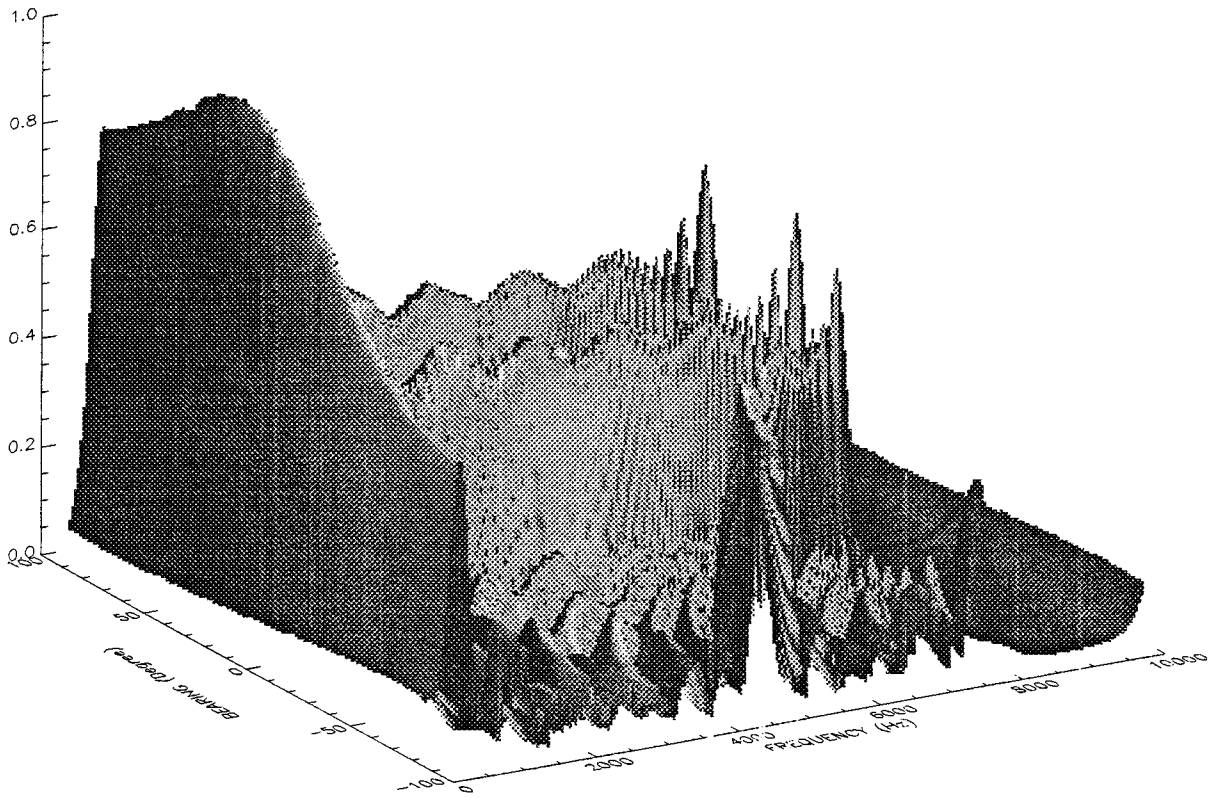
FIG. 2. Frequency domain (LFM).

equation leads to $F(\omega(1-M_r))$ as defined above using the constant velocity approximation for $r(t)$.

The range domain representations are obtained by the inverse Fourier transform. For the beamformer spectrum output

$$g_y(t,\theta,\theta_s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Y(\omega(1-M_r),\theta,\theta_s)X^*(\omega)e^{j\omega t}d\omega,$$

$$(31)$$

and for the fourth-order cumulant spectrum output,

$$g_c(t,\theta,\theta_s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} C(\omega(1-M_r),\omega_a,\omega_b,\theta,\theta_s)e^{j\omega t}d\omega.$$

$$(32)$$

These representations, $g_y(t,\theta,\theta_s)$ and $g_c(t,\theta,\theta_s)$, will be scaled to represent range in the following figures as, $range = c(t-P_t/2)$, where P_t is the transmitted pulse duration, $c = 1500$ m/s, and the initial range $r(0)$ is set to zero for the matched filter case.

The data were sampled at a sampling rate of F_s = 10 240 samples/s. Since the complex data are bandlimited so that its Fourier spectrum contains no frequencies above F_s, it can be uniquely recovered by complex values at intervals $T = 1/F_s$ from the interpolation formula described by Woodward,[30]

$$y(t) = \sum_{n=-\infty}^{\infty} y(nT)e^{j(F_o + B/2)(t-nT)}$$

$$\times (BT)\frac{\sin[\pi B(t-nT)]}{\pi B(t-nT)},$$

$$(33)$$

where $F_s = 1/T$ is greater than the bandwidth B of the transmitted signal. The above equation can be put in the same form as Woodward's[30] by letting $f_0 = F_0 + B/2$ and $W = F_s = B = 1/T$.

II. PERFECTLY REFLECTING TARGET

A. LFM transmitted waveform (1 Hz–8 kHz): Matched case

Figure 2 above shows the normalized output of the beamformer matched to the transmitted waveform in the frequency domain. The data were normalized by dividing the output by its maximum absolute value. The look angles are from $-\pi/2$ to $\pi/2$, and the band is 0 Hz to 10 kHz. Two LFM waveforms with bandwidths from 1 Hz to 8 kHz, are included, one arriving from 0° and the other from 45°. These are the only two arriavl directions that will be shown. However, other arrival directions for a given array for the most part only change the frequency at which the grating lobes start but not the basic results. In Fig. 2, the low frequencies are spatially broadened and grating lobes begin near the array design frequency. Figure 3 is the range domain representation [see Eq. (31)] of the data shown in Fig. 2. Here the
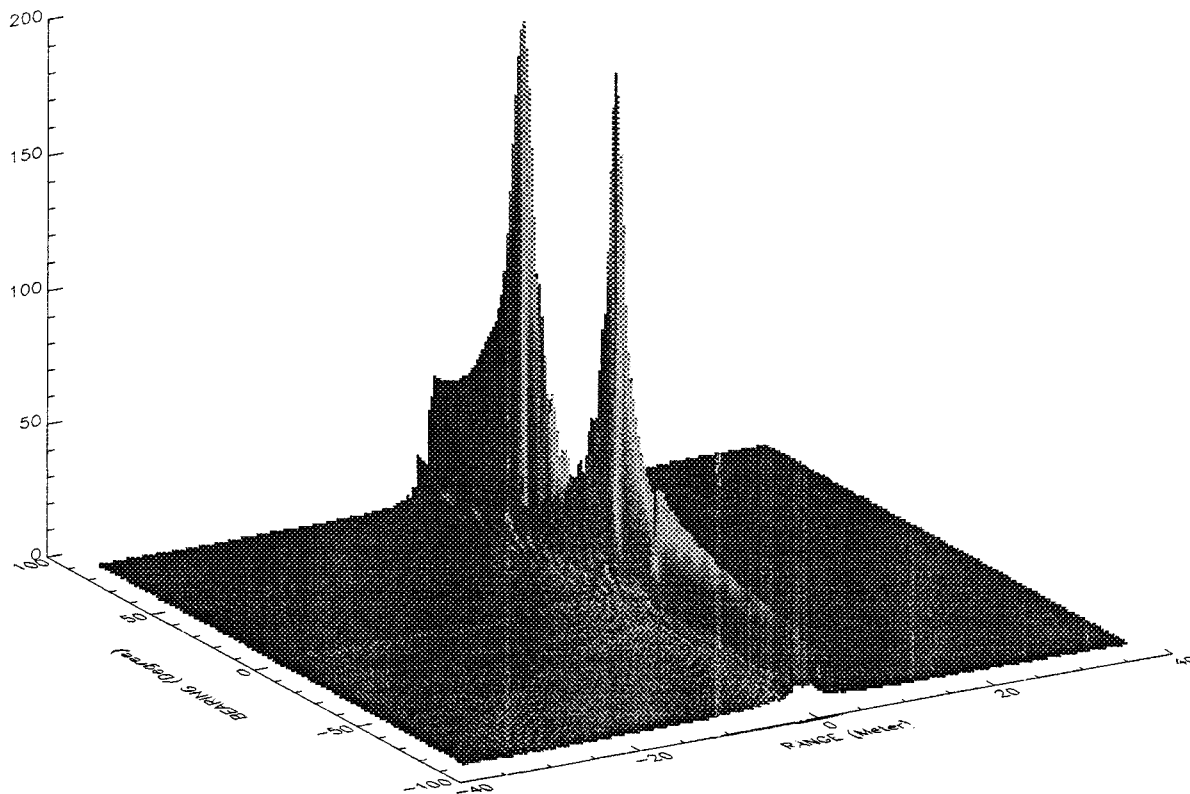
FIG. 3. Range domain (LFM).

two waveform directions are clearly depicted but the data have been spatially spread over other look directions. Figures 4 and 5 are contour plots of Figs. 2 and 3, respectively.

From the fourth-order cumulant spectrum of the beam-former output shown in Fig. 6 there is a clear improvement in the low-frequency region and a significant reduction in grating lobes compared with Fig. 2. However, the level of the fourth-order cumulant spectrum beyond the 8-kHz band in Fig. 6 is not zero. This is caused by the selection of the fixed frequencies ω_a and ω_b in the fourth-order cumulant spectrum but the level beyond the band would drop off significantly for the fourth-order moment spectrum.

The range domain results [see Eq. (32)] in Fig. 7 also show reduction of spatial spreading. The corresponding contour plots are shown in Figs. 8 and 9, respectively. There is a small peak not associated with the target directions in Figs. 7 and 9. It is also caused by the selection of the fixed frequencies, ω_a and ω_b in the fourth-order cumulant spectrum. The level of the small peak would diminish for the fourth-order moment spectrum. By using both the frequency and range domains these nontarget small peaks, when they occur, can be identified for both the fourth-order cumulant spectrum
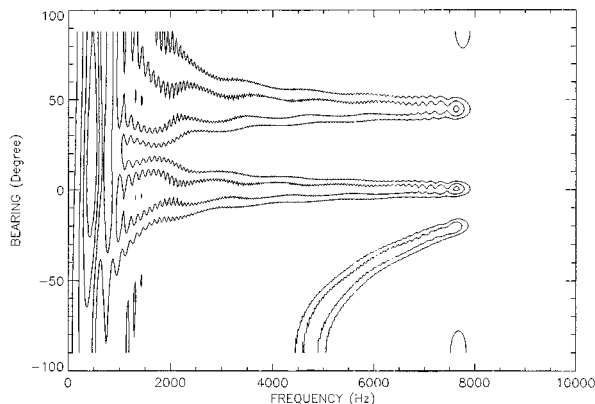
FIG. 4. Contour frequency domain (LFM).
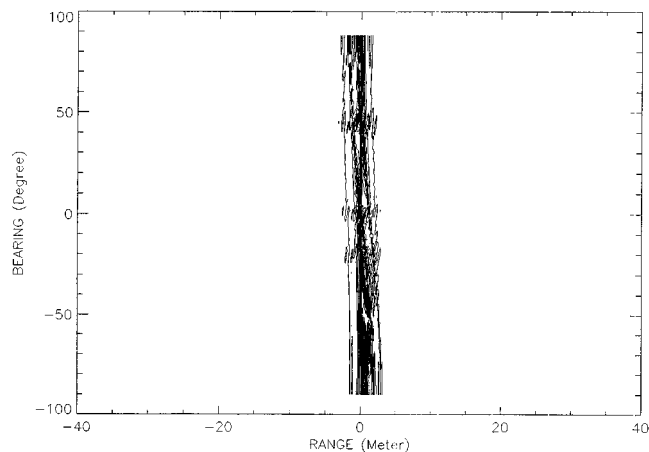
FIG. 5. Contour range domain (LFM).
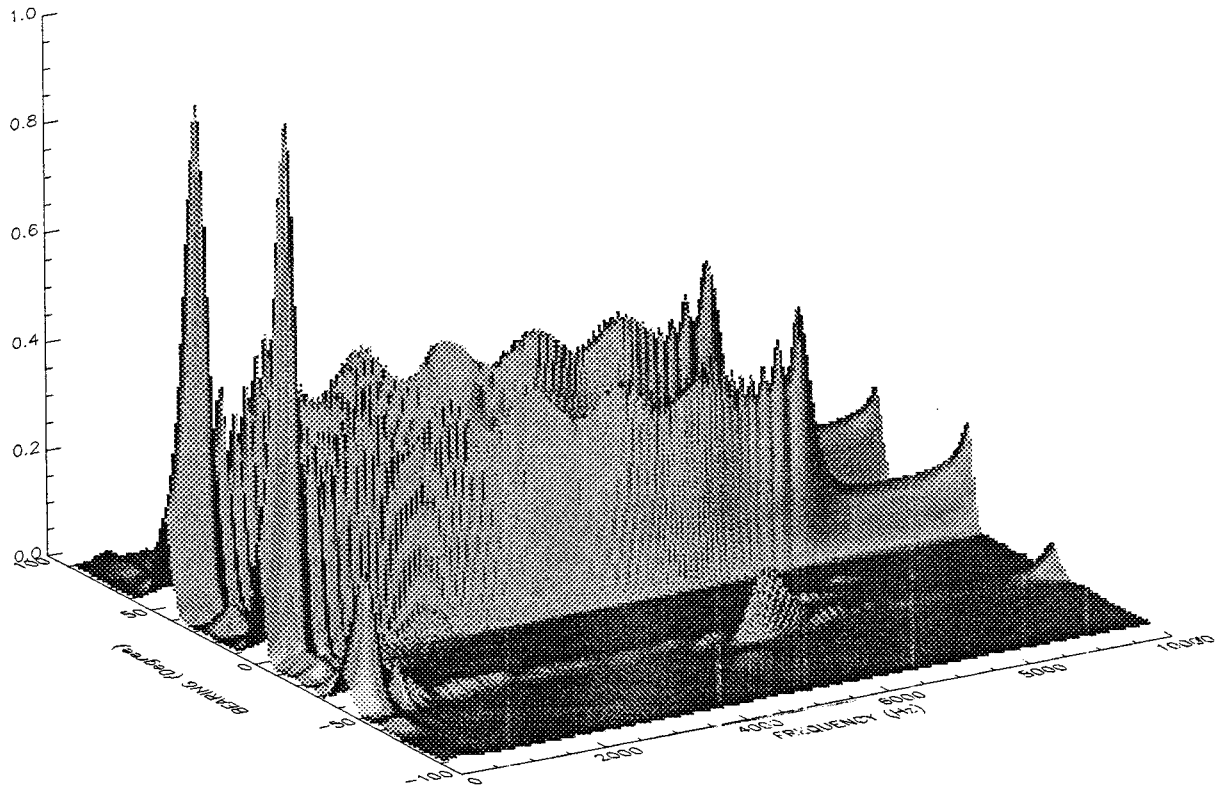
FIG. 6. Fourth-order cumulant (LFM).
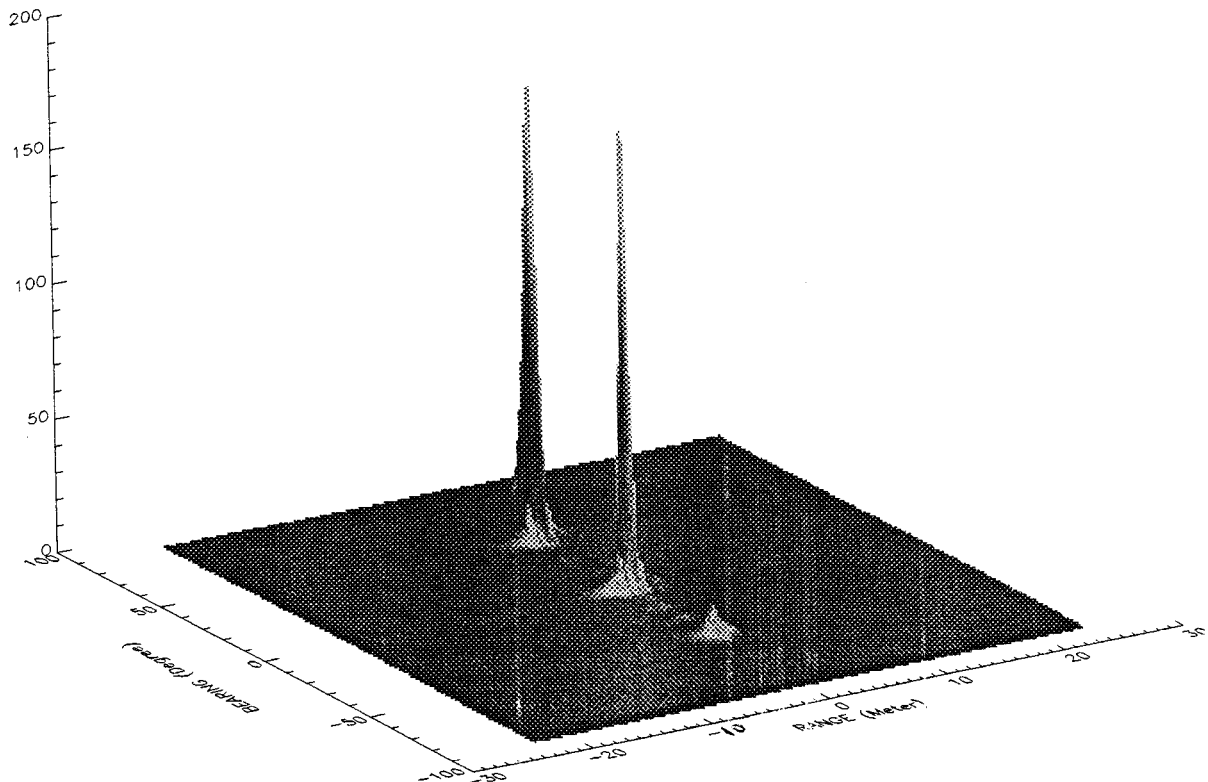
FIG. 7. Range domain (LFM).
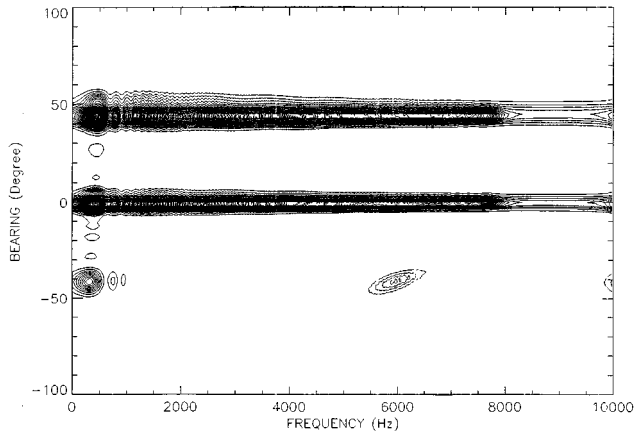
FIG. 8. Contour fourth-order cumulant (LFM).
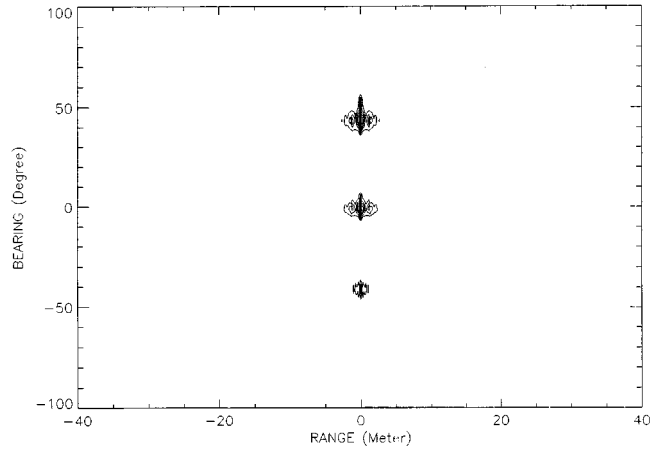
FIG. 9. Contour range domain (LFM).

and the fourth-order moment spectrum. However, nontarget peaks were not discernible in the Figures for the spherical data to be shown later.

B. LFM transmitted waveform (1 Hz–8 kHz): Mismatched case

Figures 10–13 are for the mismatched filtered beam-former output data. These are similar to the matched filtered case except the results are degraded especially for the range domain data. Because the data are mismatched,[31] i.e., matched to a zero Doppler target, there is no noticable increase in the band in Figs. 10 and 12 as would be expected for an ap-

proaching target at 28 kn. These results for the mismatched filter will now be compared to the fourth-order cumulant results.

Figures 14–17 are the corresponding plots for a 28-kn target from the fourth-order cumulant data of the beam-former output. These results do not show much change due to target motion, compared with the previous results shown in Figs. 6–9. But the small peak not associated with the target directions also occurs in Figs. 15 and 17. In addition, the band as shown in Figs. 14 and 16 has increased beyond 8 kHz as expected because the target is approaching at 28 kn. This is seen more clearly for the fourth-order moment data shown in Figs. 18 and 20.
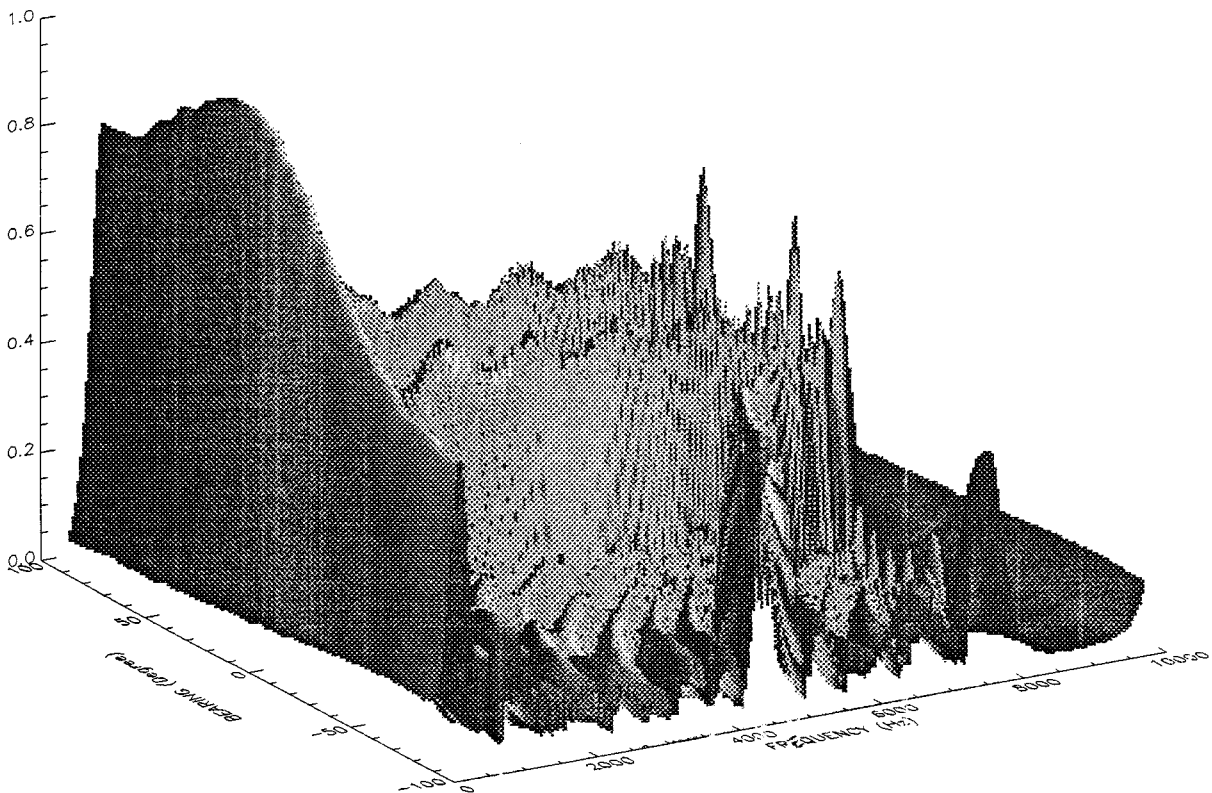
FIG. 10. Frequency domain (LFM).
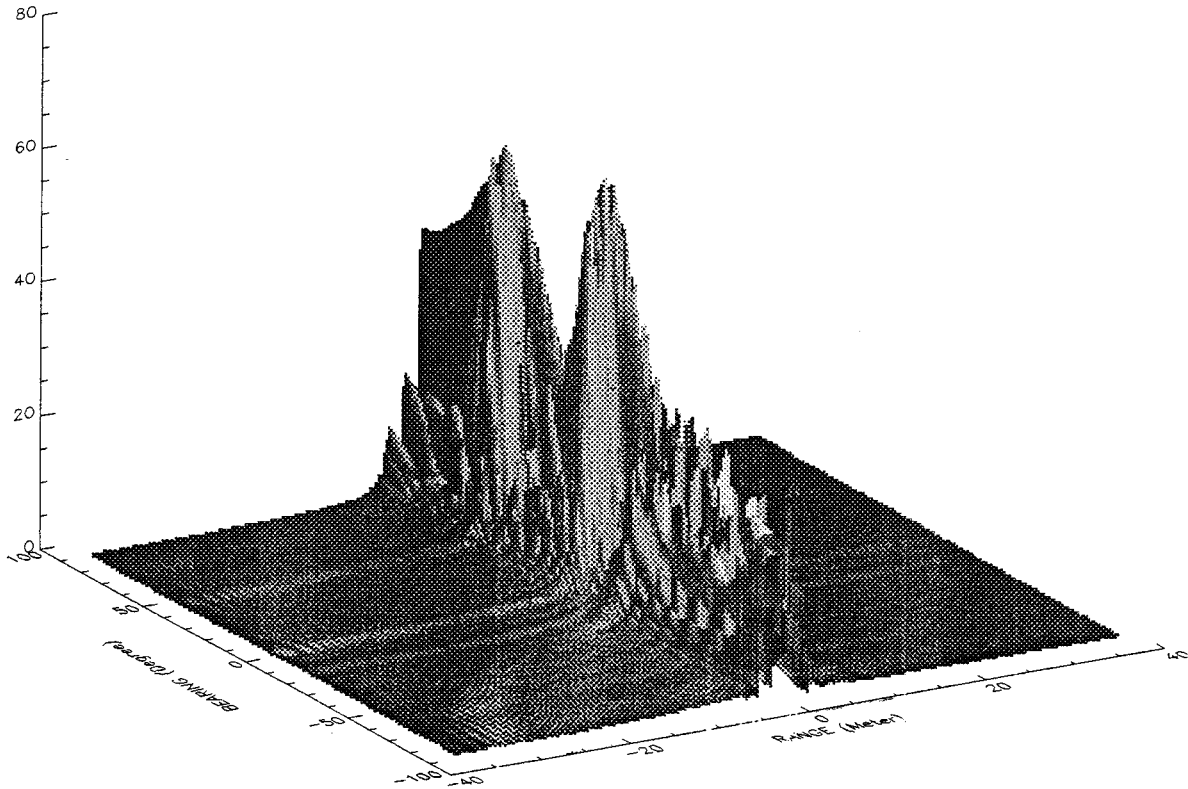
FIG. 11. Range domain (LFM).

III. SPHERICAL SHELLS

Spheres are important structures that can model underwater vehicles and other underwater structures.[28] They have also been used as a target strength reference for evaluating the sonar capabilities of dolphins.[29] Their acoustical properties have been studied extensively.[27,28] They are used here to demonstrate the improvement in extracting the spectra of spheres of different materials from propagating signals using the fourth-order cumulant spectrum of the beamformer output. Five air-filled thin spherical shells were used to demonstrate the results, consisting of a rigid sphere, a Lucite sphere, a steel sphere, an aluminum sphere, and a nickel sphere.

Because of space limitations only the results for the Lucite sphere will be shown. But the results for the five spheres mentioned above were about the same as for the Lucite sphere that will be shown. The ideal spectrum for the Lucite sphere is shown in Fig. 18 below for a 10-kHz bandwidth.

IV. SPHERICAL TARGETS

The following results represent simulated active sonar returns from the ideal spherical shells[27,28] shown above that have been convolved with a simulated transmitted LFM waveform and returned to a ten element array at a specified
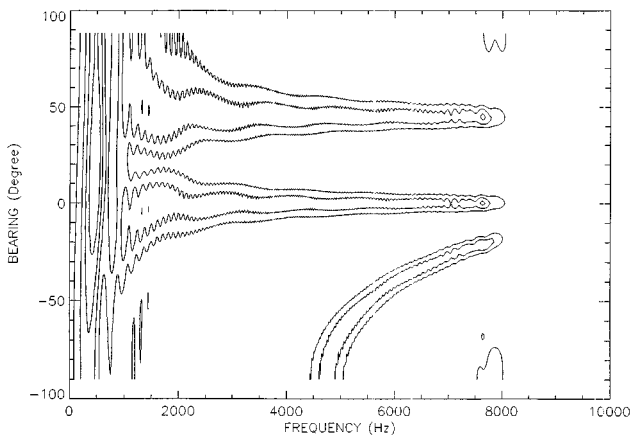
FIG. 12. Contour frequency domain (LFM).
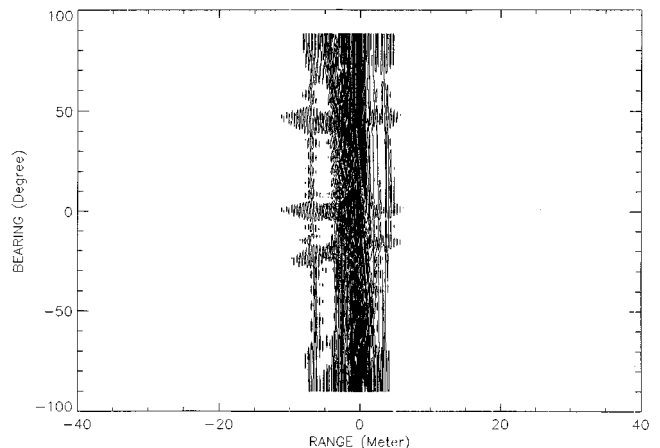
FIG. 13. Contour range domain (LFM).
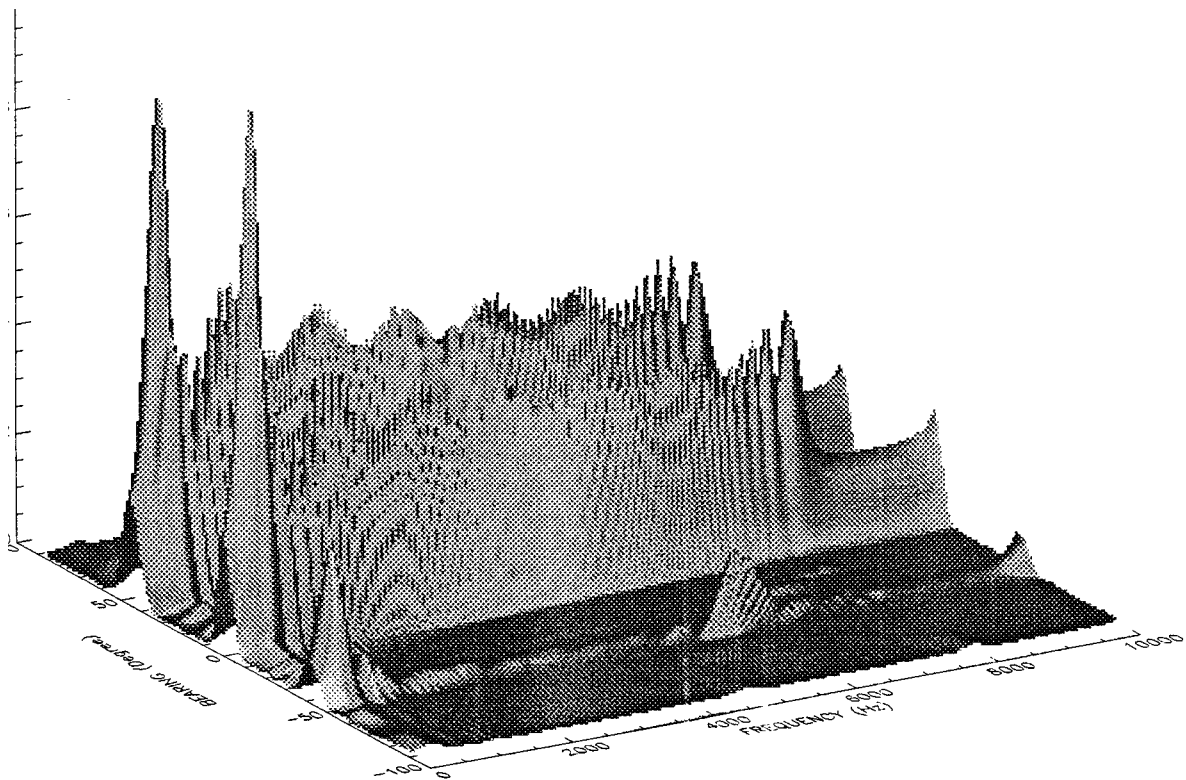
FIG. 14. Fourth-order cumulant (LFM).
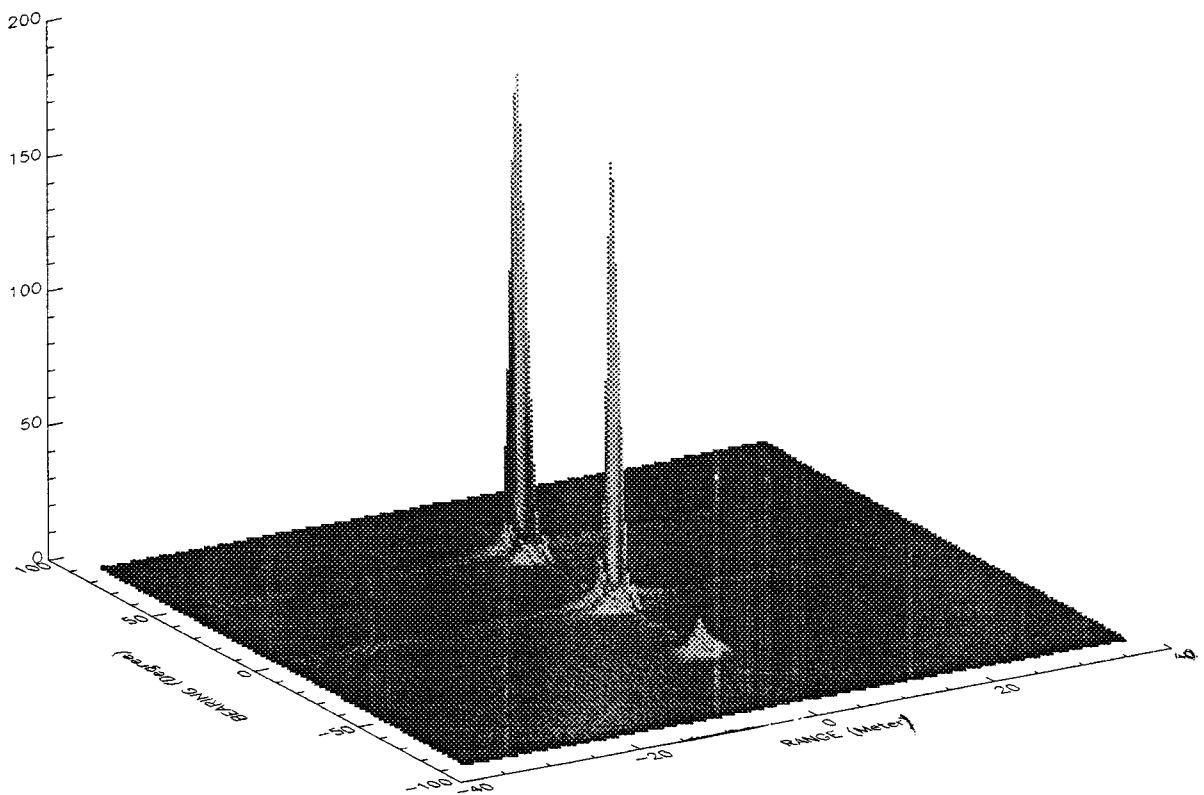
FIG. 15. Range domain (LFM).
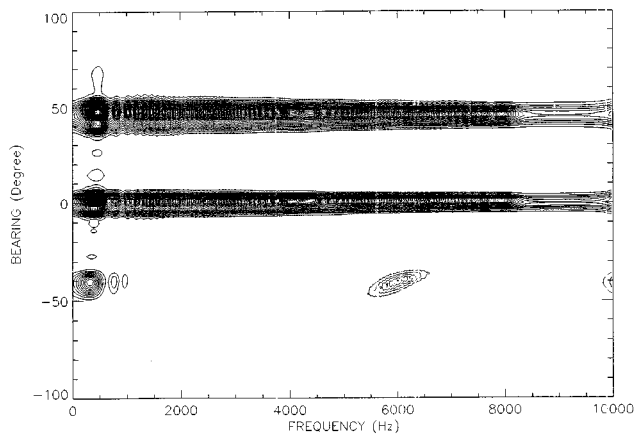
FIG. 16. Contour fourth-order cumulant (LFM).

FIG. 18. Lucite sphere.

arrival direction. Each target is either stationary or moving at 28 kn with respect to the receiving array. Since the fourth-order cumulant spectrum does not assume target velocity[15] as is the case for a matched filter[15] the two cases of a stationary target and a moving target are compared.

A. LFM transmitted waveform (1 Hz–8 kHz); target, Lucite sphere: Matched case

The characteristic resonances of a Lucite sphere[28] as shown in Fig. 19 for the spectrum of the beamformer output are spread across the spatial arrival directions over nearly the entire band. This is seen clearly in the corresponding contour plot in Fig. 21. Its time domain representation in Fig. 20 also shows spatial spreading which can be seen in the contour of Fig. 22.

For the fourth-order cumulant spectrum of the beamformer output, the characteristic resonances of a Lucite sphere are also seen in Fig. 23 but only at the arrival directions. Its contour plot depicts this significant improvement in Fig. 25. The time domain representation in Fig. 24 and the corresponding contour in Fig. 26 clearly show the arrival directions for these Lucite spheres without spatial spreading. The data along the arrival directions in Figs. 24 and 26 are related to the autocorrelation of the impulse response of the lucite sphere since two frequencies of the fourth-order cumu-

lant spectrum defined in Eq. (32) are fixed. No such autocorrelation relationship can be descerned from the matched filtered data. To summarize, the overall improvements are the wideband spatial resolution and enhanced target characteristics obtained with the fourth-order cumulant spectrum and its range domain representation of the beamformer output compared with the matched filtered spectrum and its time domain representation of the beamformer output.
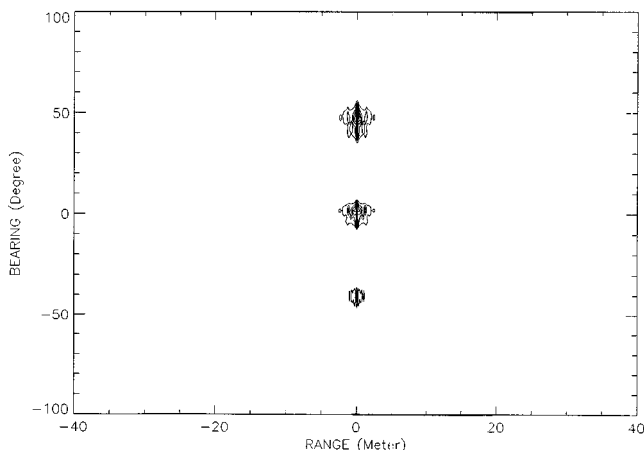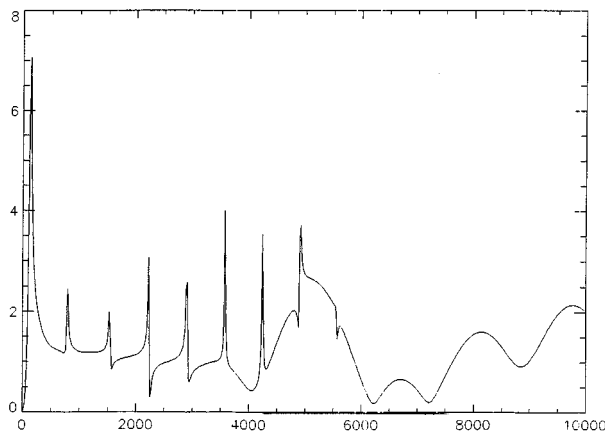
B. LFM transmitted waveform (1 Hz–8 kHz); target, Lucite sphere: Mismatched case

In the mismatched case the characteristic resonances of a Lucite sphere[28] as shown in Fig. 27 and its contour in Fig. 29 for the spectrum of the beamformer output are also spread across the spatial arrival directions over nearly the entire band as in the matched case. However, its range domain representation in Fig. 28 and its contour in Fig. 30 show significantly more spatial spreading compared with the matched case.

For the fourth-order cumulant spectrum of the beamformer output, for the 28-kn target, the characteristic lucite resonances are present along the arrival directions as shown in Fig. 31 and in its contour in Fig. 33. The range domain representation in Fig. 32 and its contour in Fig. 34 depict clear arrival directions with the autocorrelation related data of the impulse response of the Lucite sphere.

V. CONCLUSIONS

Since sonar arrays play such an important role in underwater acoustic applications a fourth-order cumulant spectrum processing method that extends the usefullness of uniformly distributed sensor arrays in extracting information from broadband signals, compared to the array's design frequency, was developed. Array performance degradation from grating lobes and beam broadening were defined by a theorem and its corollary. Another theorem and its corollary were stated and proven for the fourth-order cumulant spectrum of the beamformer output that explained why the effects of broadband signals could be suppressed. It was demonstrated by simulating an active sonar system and spherical targets that the deleterious effects of operating a line array well beyond its designed frequency could be reduced by properly choos-

FIG. 17. Contour range domain (LFM).
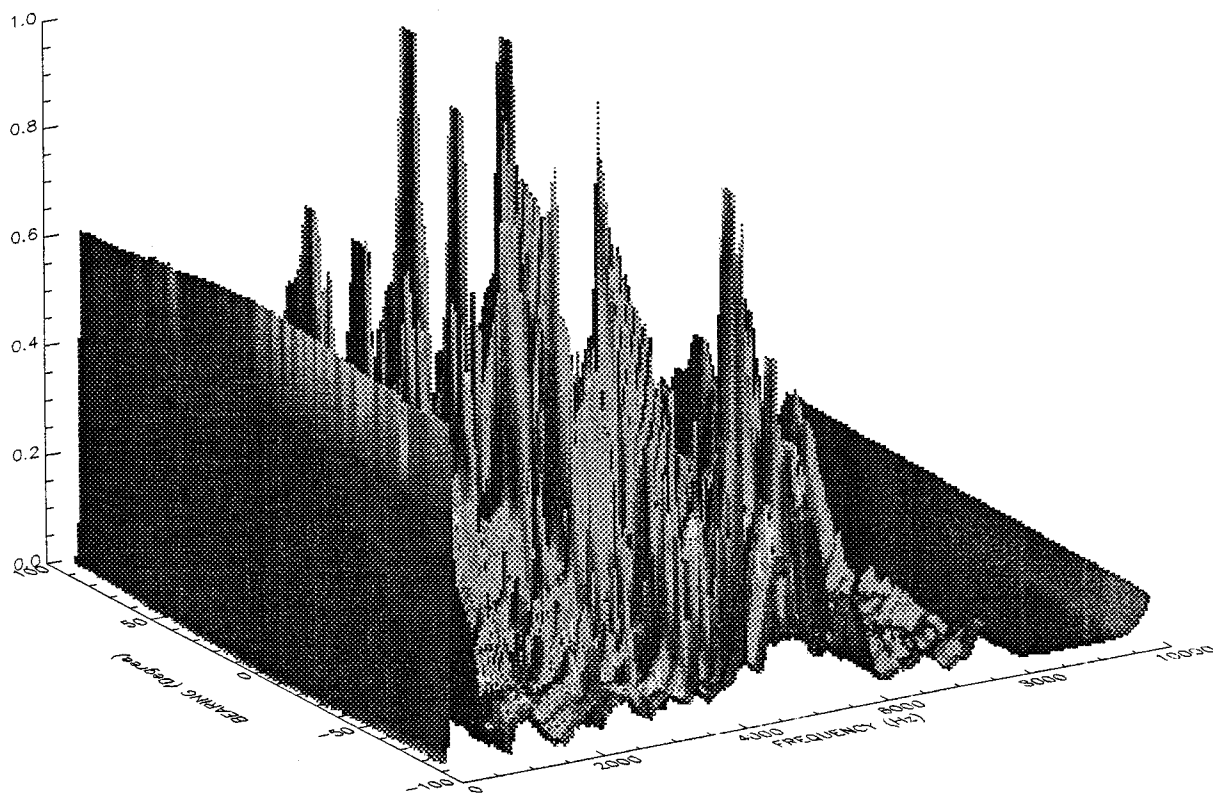
FIG. 19. Frequency domain (Lucite).
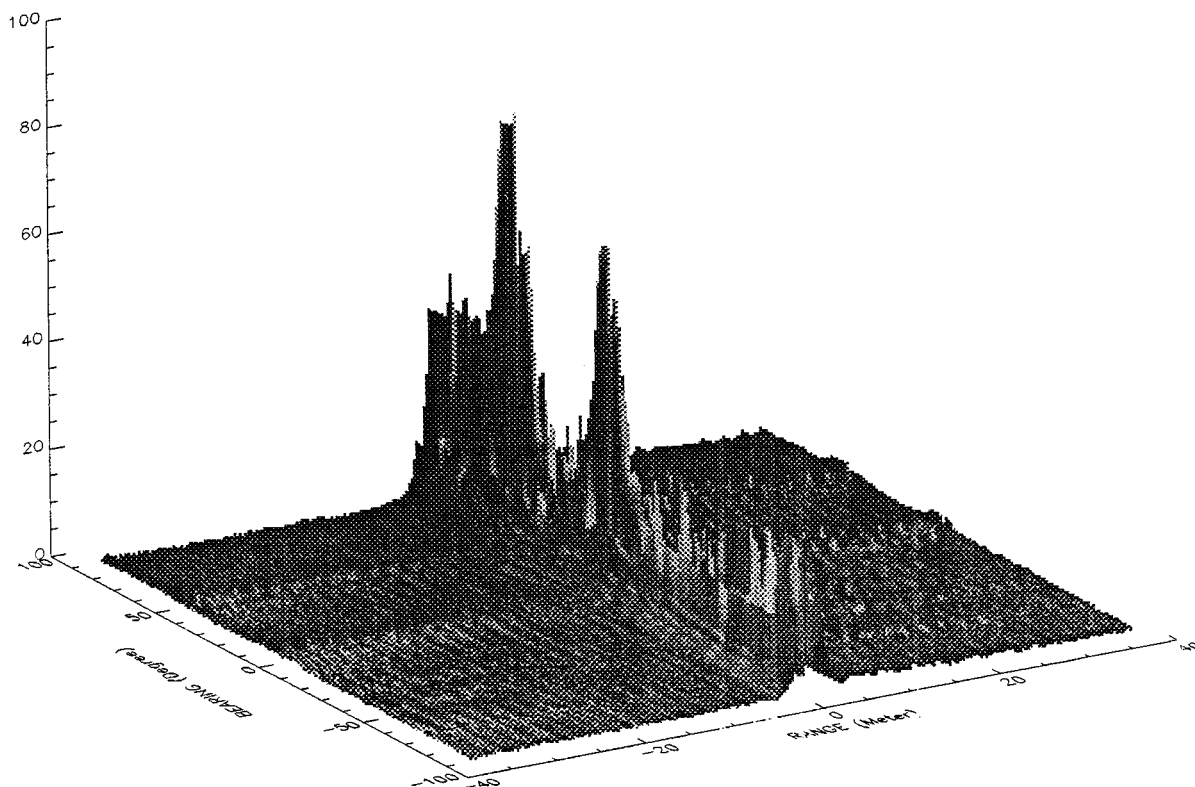
FIG. 20. Range domain (Lucite).
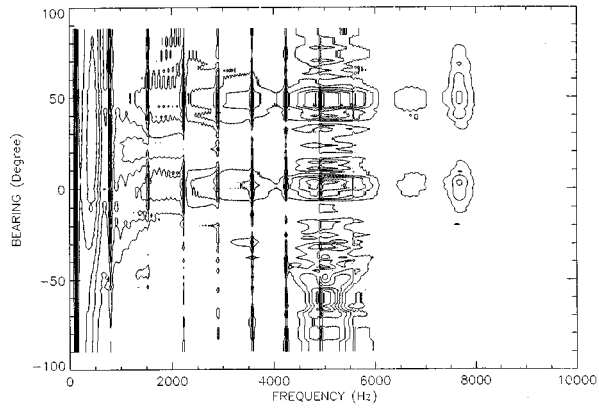
FIG. 21. Contour frequency domain (Lucite).
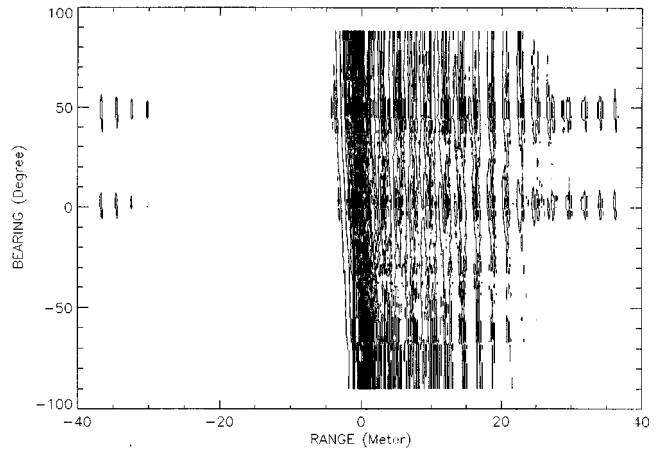
FIG. 22. Contour range domain (Lucite).
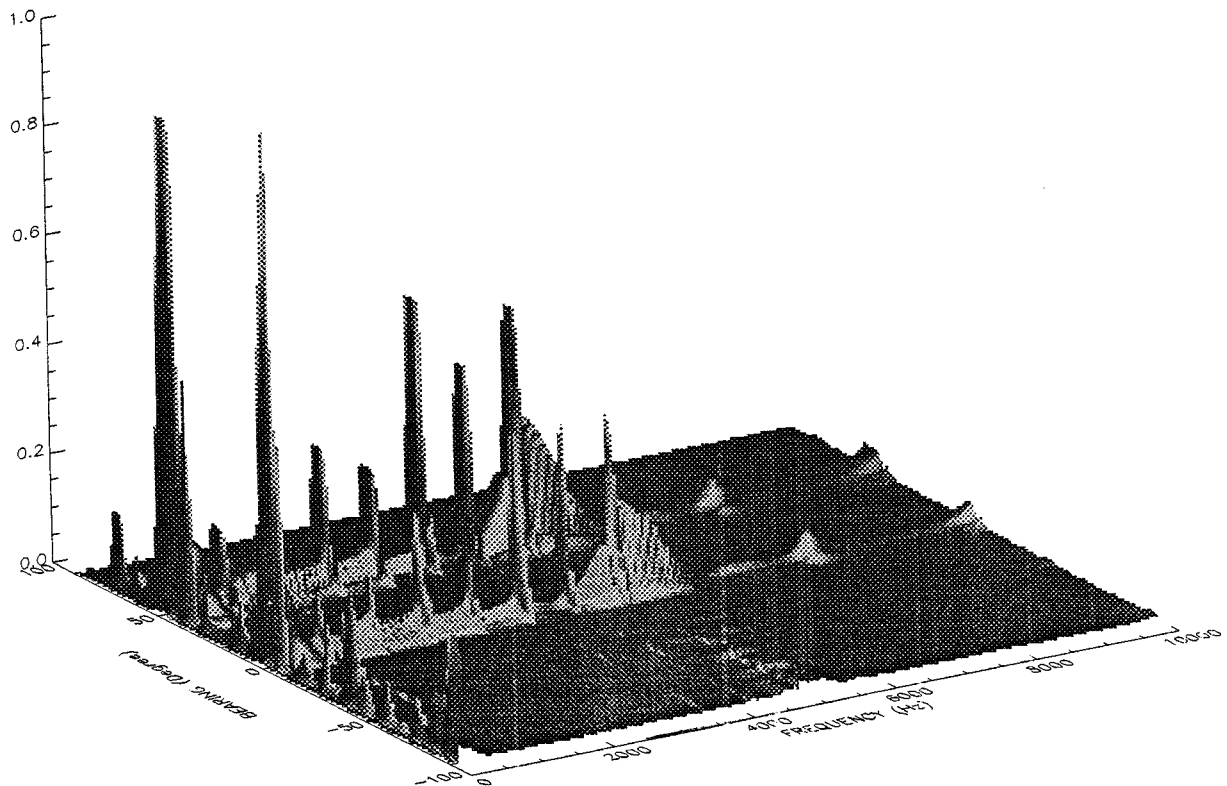
FIG. 23. Fourth-order cumulant (Lucite).
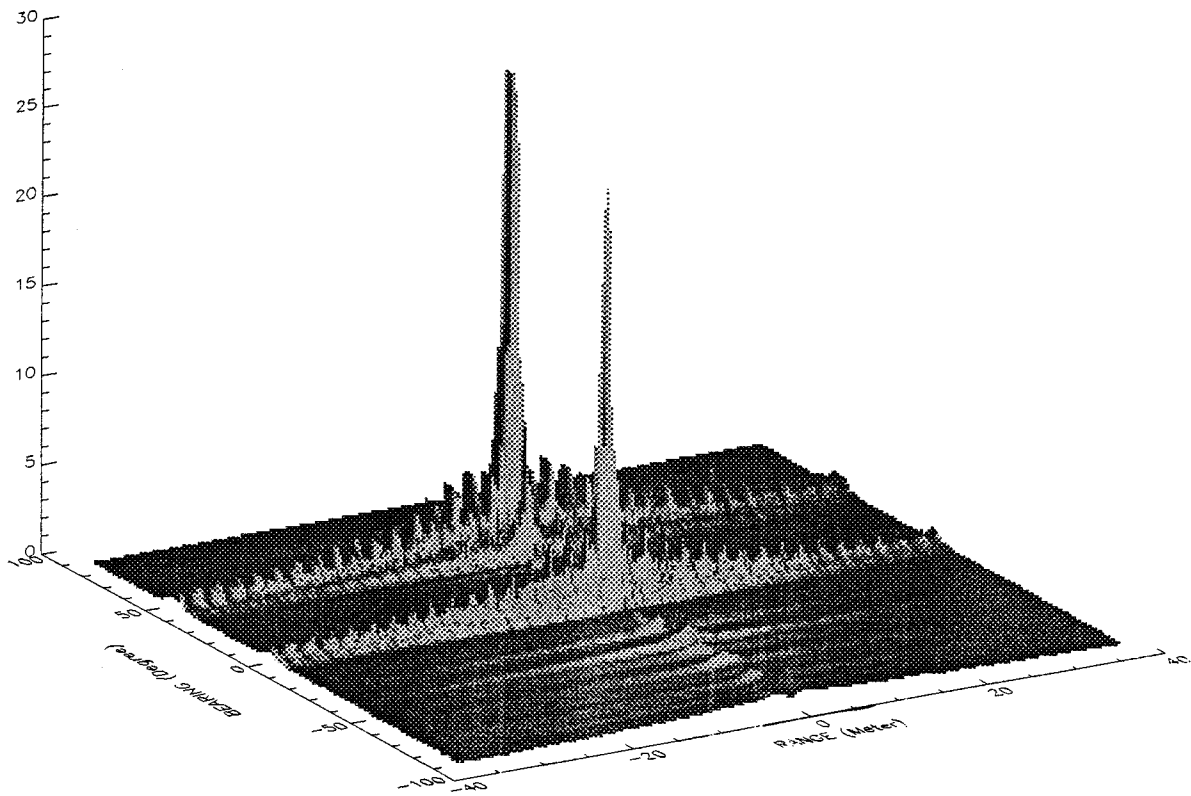
FIG. 24. Range domain (Lucite).
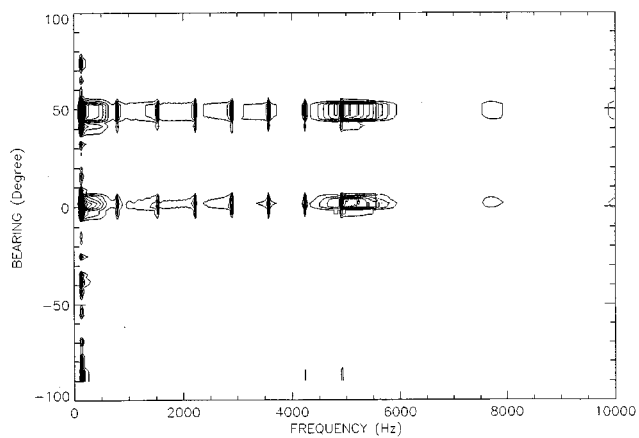
FIG. 25. Contour fourth-order cumulant (Lucite).
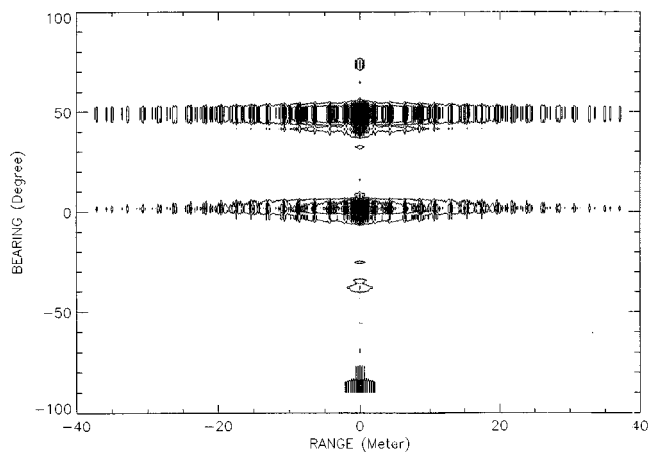
FIG. 26. Contour range domain (Lucite).
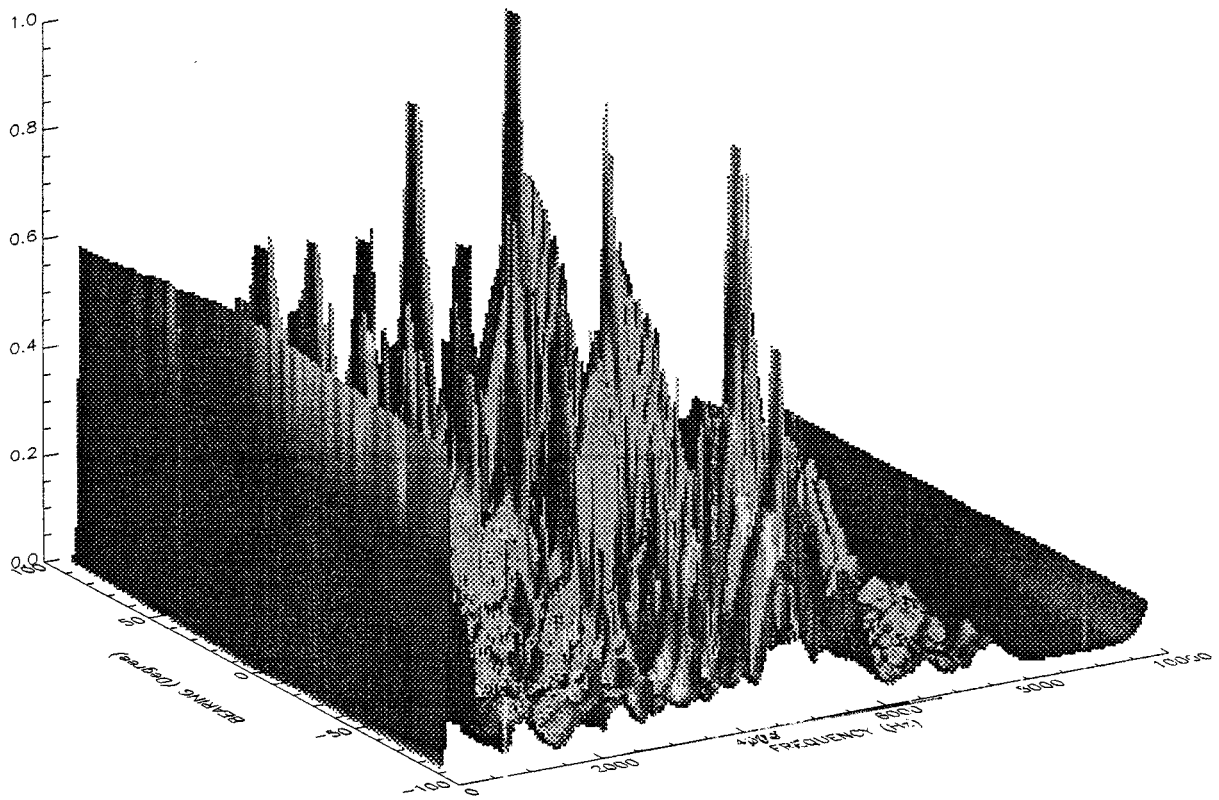
FIG. 27. Frequency domain (Lucite).
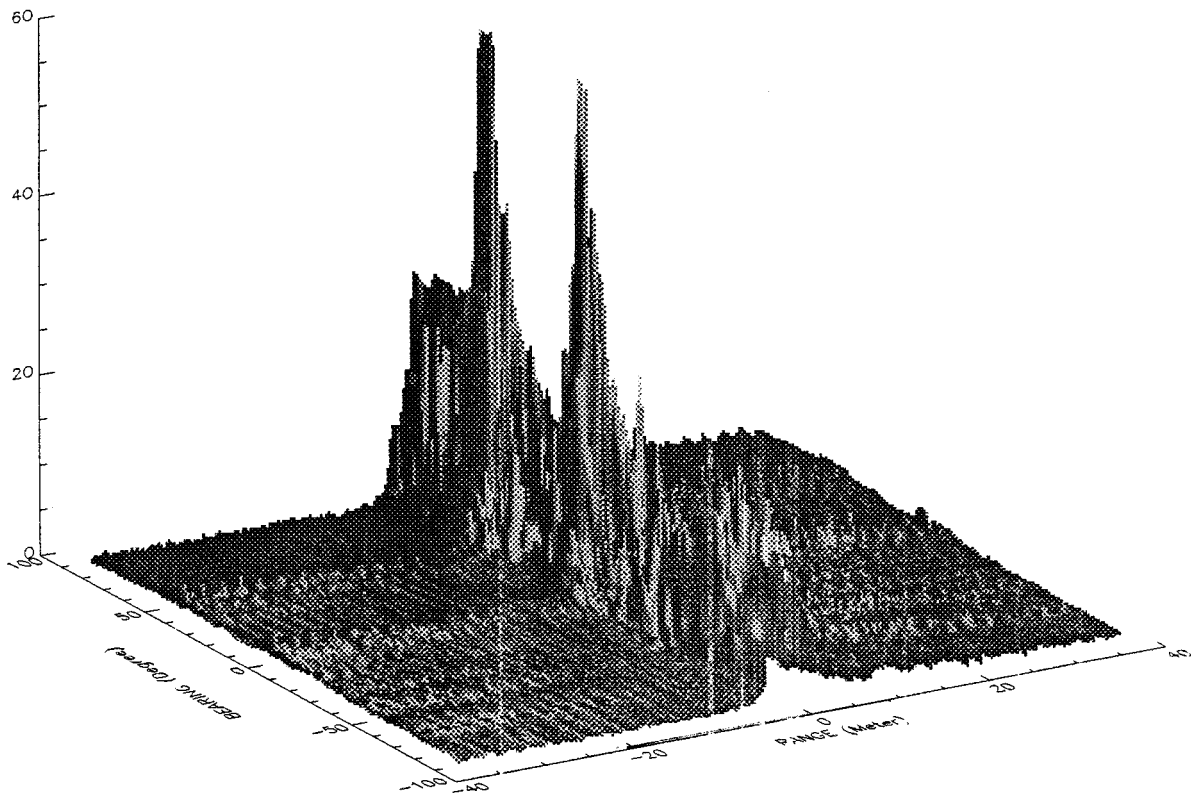
FIG. 28. Range domain (Lucite).
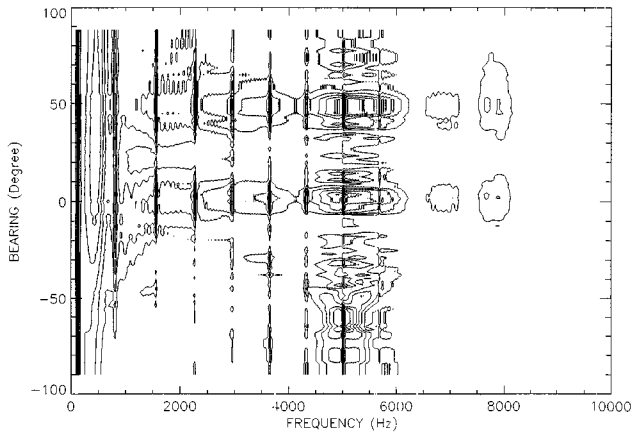
FIG. 29. Contour frequency domain (Lucite).
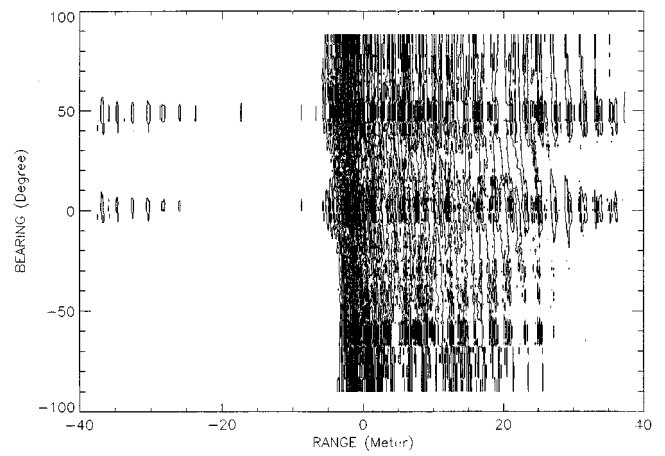
FIG. 30. Contour range domain (Lucite).

ing frequencies of the fourth-order cumulant spectrum of the beamformer output. Five types of spheres were simulated, convolved with a linear frequency modulated waveform and given spatial directivity in order to compare the results of a beamformer's spectrum and its fourth-order cumulant spectrum. The results demonstrated that the fourth-order cumulant spectrum of a beamformer output and its range domain representation, with two of its three frequencies properly chosen and fixed, improved performance rather dramatically

in most cases examined, in terms of reduced spatial spreading, increased main lobe resolution at low frequencies, and reduced grating lobes at high frequencies compared with the spectrum of the beamformer output and its range domain representation. It was also theoretically shown that the phase of an input signal could be extracted at the beamformer output without distortion by the fourth-order moment spectrum over a wide range of conditions that normally would cause distortion of the phase.
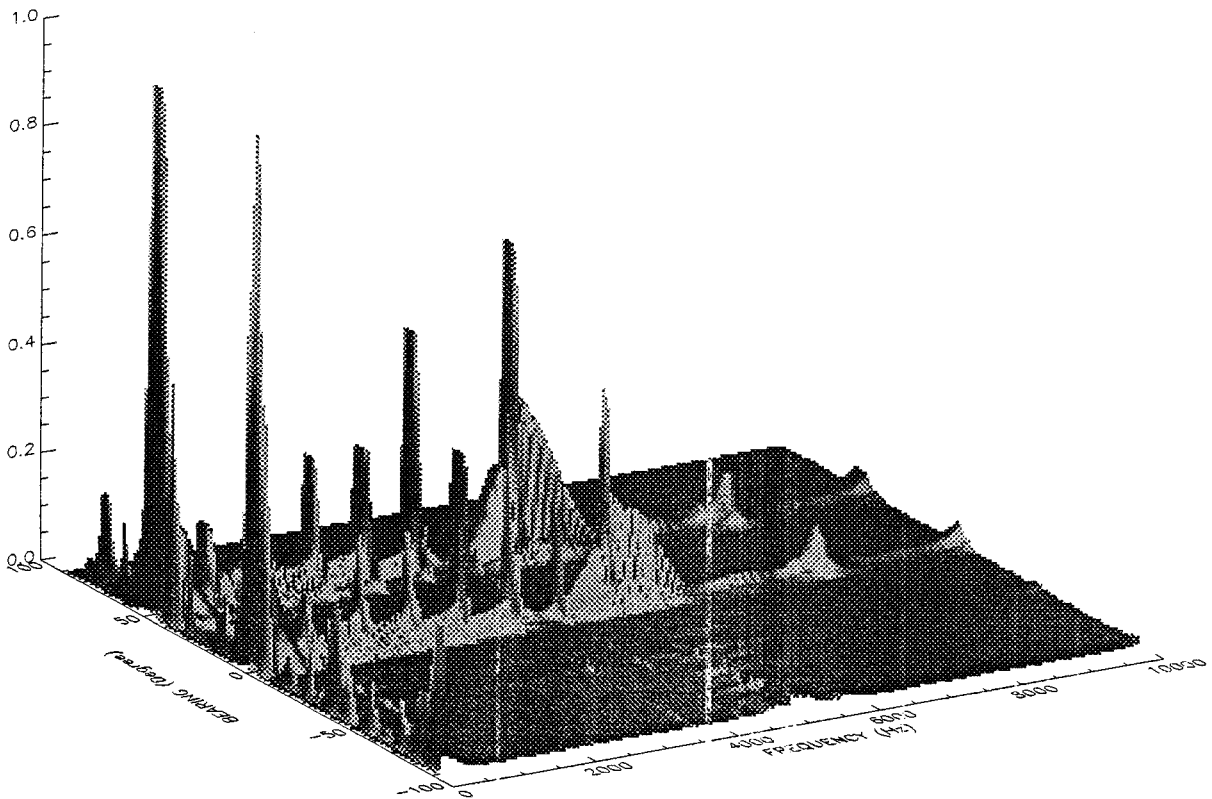
FIG. 31. Fourth-order cumulant (Lucite).
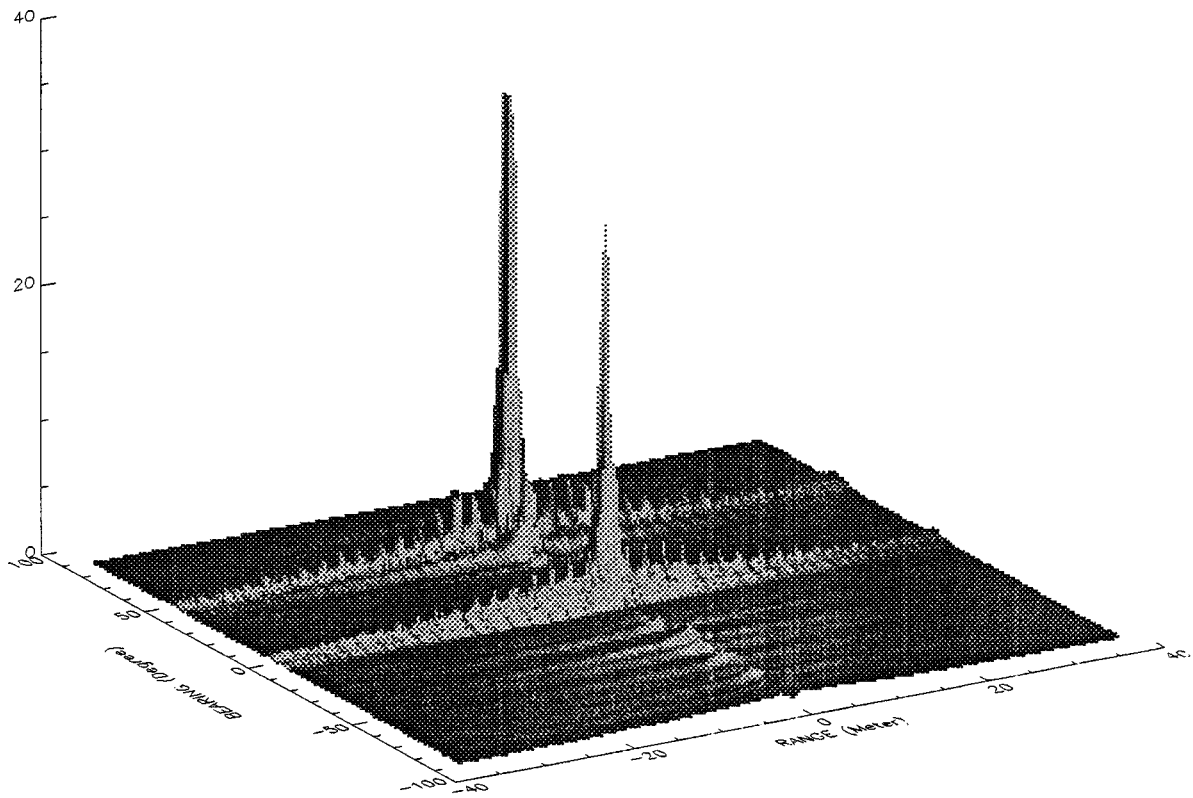
FIG. 32. Range domain (Lucite).
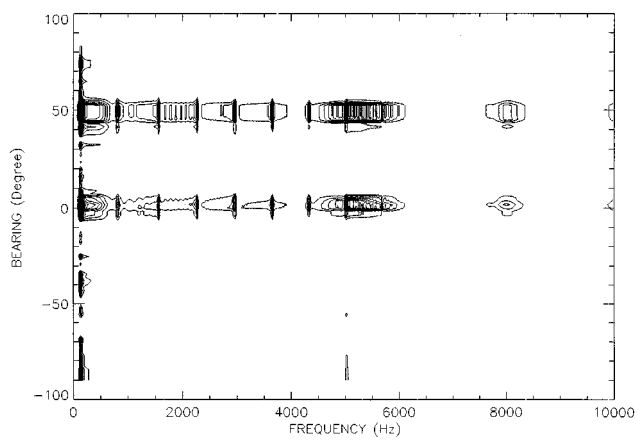
FIG. 33. Contour fourth-order cumulant (Lucite).
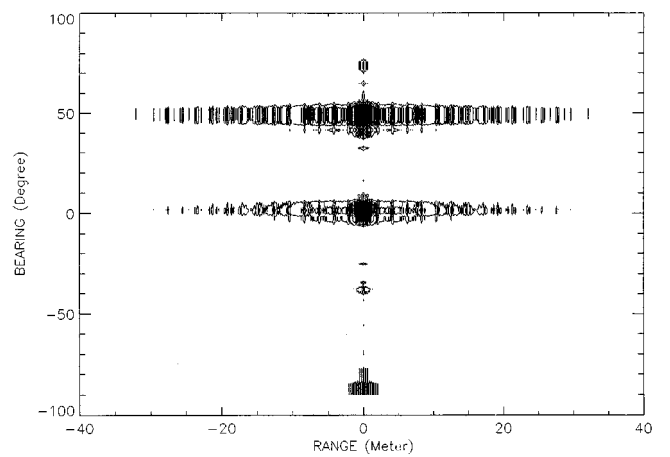
FIG. 34. Contour range domain (Lucite).

APPENDIX

The fourth-order cumulant spectrum at the output of a beamformer for N_t targets is given by

$$C_4(\omega_1,\omega_2,\omega_3,\Theta,\theta_s)$$

$$= \sum_{i=1}^{N_t} \{G_i(\omega_1)G_i(\omega_2)G_i(\omega_3)G_i^*(\omega_4) - |G_i(\omega_1)|^2|G_i(\omega_2)|^2 - |G_i(\omega_1)|^2|G_i(\omega_3)|^2 - |G_i(\omega_2)|^2|G_i(\omega_3)|^2\}$$

$$+ \sum_{i_1}^{N_t-1}\sum_{i_2=i_1+1}^{N_t} \{[G_{i_1}(\omega_1)G_{i_1}(\omega_2)G_{i_2}(\omega_3)G_{i_2}^*(\omega_4) + G_{i_2}(\omega_1)G_{i_2}(\omega_2)G_{i_1}(\omega_3)G_{i_1}^*(\omega_4)]$$

$$- [|G_{i_1}(\omega_1)|^2|G_{i_2}(\omega_2)|^2 + |G_{i_2}(\omega_1)|^2|G_{i_1}(\omega_2)|^2] - [|G_{i_1}(\omega_1)|^2|G_{i_2}(\omega_3)|^2 + |G_{i_2}(\omega_1)|^2|G_{i_1}(\omega_3)|^2]$$

$$- [|G_{i_1}(\omega_2)|^2|G_{i_2}(\omega_3)|^2 + |G_{i_2}(\omega_2)|^2|G_{i_1}(\omega_3)|^2]\}$$

$$+ \sum_{i_1}^{N_t-1}\sum_{i_2=i_1+1}^{N_t} \{G_{i_1}(\omega_3)G_{i_1}^*(\omega_4)[G_{i_1}(\omega_1)G_{i_2}(\omega_2) + G_{i_2}(\omega_1)G_{i_1}(\omega_2)] + G_{i_2}(\omega_3)G_{i_2}^*(\omega_4)[G_{i_1}(\omega_1)G_{i_2}(\omega_2)$$

$$+ G_{i_2}(\omega_1)G_{i_1}(\omega_2)] + G_{i_1}(\omega_1)G_{i_1}(\omega_2)[G_{i_1}(\omega_3)G_{i_2}^*(\omega_4) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_4)] + G_{i_2}(\omega_1)G_{i_2}(\omega_2)[G_{i_1}(\omega_3)G_{i_2}^*(\omega_4)$$

$$+ G_{i_2}(\omega_3)G_{i_1}^*(\omega_4)] + [G_{i_1}(\omega_1)G_{i_2}^*(\omega_2) + G_{i_2}(\omega_1)G_{i_1}^*(\omega_2)][G_{i_1}(\omega_3)G_{i_2}^*(\omega_4) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_4)]$$

$$- (|G_{i_1}(\omega_2)|^2[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1) + G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)] + |G_{i_2}(\omega_2)|^2[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1) + G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)]$$

$$+ |G_{i_1}(\omega_1)|^2[G_{i_1}(\omega_2)G_{i_2}^*(\omega_2) + G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)] + |G_{i_2}(\omega_1)|^2[G_{i_1}(\omega_2)G_{i_2}^*(\omega_2) + G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)]$$

$$+ [G_{i_1}(\omega_1)G_{i_2}^*(\omega_1) + G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_1}(\omega_2)G_{i_2}^*(\omega_2) + G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)])$$

$$- (|G_{i_1}(\omega_3)|^2[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1) + G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)] + |G_{i_2}(\omega_3)|^2[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1) + G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)]$$

$$+ |G_{i_1}(\omega_1)|^2[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)] + |G_{i_2}(\omega_1)|^2[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)]$$

$$+ [G_{i_1}(\omega_1)G_{i_2}^*(\omega_1) + G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_1}(\omega_3)G_{i_2}^*(\omega_3) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)])$$

$$- (|G_{i_1}(\omega_3)|^2[G_{i_1}(\omega_2)G_{i_2}^*(\omega_2) + G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)] + |G_{i_2}(\omega_3)|^2[G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)$$

$$+ G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)] + |G_{i_1}(\omega_2)|^2[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)$$

$$+ G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)] + |G_{i_2}(\omega_2)|^2[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)]$$

$$+ [G_{i_1}(\omega_2)G_{i_2}^*(\omega_2) + G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)][G_{i_1}(\omega_3)G_{i_2}^*(\omega_3) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)])\}$$

$$+ \sum_{i_1=1}^{N_t-2}\sum_{i_2=i_1+1}^{N_t-1}\sum_{i_3=i_2+1}^{N_t} \{G_{i_1}(\omega_3)G_{i_1}^*(\omega_4)[G_{i_2}(\omega_1)G_{i_3}(\omega_2) + G_{i_3}(\omega_1)G_{i_2}(\omega_2)]$$

$$+ G_{i_2}(\omega_3)G_{i_2}^*(\omega_4)[G_{i_1}(\omega_1)G_{i_3}(\omega_2) + G_{i_3}(\omega_1)G_{i_1}(\omega_2)] + G_{i_3}(\omega_3)G_{i_3}^*(\omega_4)[G_{i_1}(\omega_1)G_{i_2}(\omega_2) + G_{i_2}(\omega_1)G_{i_1}(\omega_2)]$$

$$+ G_{i_1}(\omega_1)G_{i_1}(\omega_2)[G_{i_2}(\omega_3)G_{i_3}^*(\omega_4) + G_{i_3}(\omega_3)G_{i_2}^*(\omega_4)]$$

$$+ G_{i_2}(\omega_1)G_{i_2}(\omega_2)[G_{i_1}(\omega_3)G_{i_3}^*(\omega_4) + G_{i_3}(\omega_3)G_{i_1}^*(\omega_4)] + G_{i_3}(\omega_1)G_{i_3}(\omega_2)$$

$$\times [G_{i_1}(\omega_3)G_{i_2}^*(\omega_4) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_4)] + [G_{i_1}(\omega_1)G_{i_2}(\omega_2) + G_{i_2}(\omega_1)G_{i_1}(\omega_2)][G_{i_2}(\omega_3)G_{i_3}^*(\omega_4) + G_{i_3}(\omega_3)G_{i_2}^*(\omega_4)]$$

$$+ [G_{i_1}(\omega_1)G_{i_3}(\omega_2) + G_{i_3}(\omega_1)G_{i_1}(\omega_2)][G_{i_3}(\omega_3)G_{i_2}^*(\omega_4) + G_{i_2}(\omega_3)G_{i_3}^*(\omega_4)] + [G_{i_2}(\omega_1)G_{i_3}(\omega_2) + G_{i_3}(\omega_1)G_{i_2}(\omega_2)]$$

$$\times [G_{i_3}(\omega_3)G_{i_1}^*(\omega_4) + G_{i_1}(\omega_3)G_{i_3}^*(\omega_4)]$$

$$+ [G_{i_3}(\omega_1)G_{i_1}(\omega_2) + G_{i_1}(\omega_1)G_{i_3}(\omega_2)][G_{i_1}(\omega_3)G_{i_2}^*(\omega_4) + G_{i_2}(\omega_3)G_{i_1}^*(\omega_4)] + [G_{i_3}(\omega_1)G_{i_2}(\omega_2) + G_{i_2}(\omega_1)G_{i_3}(\omega_2)]$$

$$\times [G_{i_2}(\omega_3)G_{i_1}^*(\omega_4) + G_{i_1}(\omega_3)G_{i_2}^*(\omega_4)] + [G_{i_2}(\omega_1)G_{i_1}(\omega_2) + G_{i_1}(\omega_1)G_{i_2}(\omega_2)]$$

$$\times [G_{i_1}(\omega_3)G_{i_3}^*(\omega_4) + G_{i_3}(\omega_3)G_{i_1}^*(\omega_4)] - (|G_{i_1}(\omega_2)|^2[G_{i_2}(\omega_1)G_{i_3}^*(\omega_1) + G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)]$$

$$+|G_{i_2}(\omega_2)|^2[G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)]+|G_{i_3}(\omega_2)|^2[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)+G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)]$$

$$+|G_{i_1}(\omega_1)|^2[G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)]$$

$$+|G_{i_2}(\omega_1)|^2[G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_1}^*(\omega_2)]+|G_{i_3}(\omega_1)|^2[G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)]$$

$$+[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)\quad+G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)]$$

$$+[G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)]$$

$$+[G_{i_2}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)][G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)]+[G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)$$

$$+G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)][G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)]+[G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)$$

$$+G_{i_2}(\omega_1)G_{i_3}^*(\omega_1)][G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)+G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)]$$

$$+[G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)+G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)][G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_1}^*(\omega_2)])$$

$$-(|G_{i_1}(\omega_3)|^2[G_{i_2}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)]+|G_{i_2}(\omega_3)|^2[G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)]$$

$$+|G_{i_3}(\omega_3)|^2[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)+G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)]+|G_{i_1}(\omega_1)|^2[G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)]$$

$$+|G_{i_2}(\omega_1)|^2[G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_1}^*(\omega_3)]$$

$$+|G_{i_3}(\omega_1)|^2[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)]+[G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)+G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)$$

$$+G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)]+[G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)]$$

$$+[G_{i_2}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)][G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)]+[G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)+G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)]$$

$$\times[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)]+[G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)+G_{i_2}(\omega_1)G_{i_3}^*(\omega_1)][G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)+G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)]$$

$$+[G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)+G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)][G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_1}^*(\omega_3)])$$

$$-(|G_{i_1}(\omega_3)|^2[G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)]$$

$$+|G_{i_2}(\omega_3)|^2[G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_1}^*(\omega_2)]+|G_{i_3}(\omega_3)|^2[G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)]$$

$$+|G_{i_1}(\omega_2)|^2[G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)]$$

$$+|G_{i_2}(\omega_2)|^2[G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_1}^*(\omega_3)]+|G_{i_3}(\omega_2)|^2[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)]$$

$$+[G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)\quad+G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)][G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)]$$

$$+[G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_1}^*(\omega_2)][G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)]+[G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)]$$

$$\times[G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)]+[G_{i_3}(\omega_2)G_{i_1}^*(\omega_2)+G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)][G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)]$$

$$+[G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)][G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)+G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)]+[G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)+G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)]$$

$$\times[G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_1}^*(\omega_3)])\}$$

$$+\sum_{i_1=1}^{N_t-3}\sum_{i_2=i_1+1}^{N_t-2}\sum_{i_3=i_2+1}^{N_t-1}\sum_{i_4=i_3+1}^{N_t}\{[G_{i_1}(\omega_1)G_{i_2}(\omega_2)+G_{i_2}(\omega_2)G_{i_1}(\omega_2)][G_{i_3}(\omega_3)G_{i_4}^*(\omega_4)+G_{i_4}(\omega_3)G_{i_3}^*(\omega_4)]$$

$$+[G_{i_1}(\omega_1)G_{i_3}(\omega_2)+G_{i_3}(\omega_1)G_{i_1}(\omega_2)][G_{i_2}(\omega_3)G_{i_4}^*(\omega_4)+G_{i_4}(\omega_3)G_{i_2}^*(\omega_4)]+[G_{i_1}(\omega_1)G_{i_4}(\omega_2)+G_{i_4}(\omega_1)G_{i_1}(\omega_2)]$$

$$\times[G_{i_1}(\omega_3)G_{i_3}^*(\omega_4)+G_{i_3}(\omega_3)G_{i_2}^*(\omega_4)]+[G_{i_2}(\omega_1)G_{i_3}(\omega_2)+G_{i_3}(\omega_1)G_{i_2}(\omega_2)][G_{i_1}(\omega_3)G_{i_4}^*(\omega_4)+G_{i_4}(\omega_3)G_{i_3}^*(\omega_4)]$$

$$+[G_{i_2}(\omega_1)G_{i_4}(\omega_2)+G_{i_4}(\omega_1)G_{i_2}(\omega_2)][G_{i_1}(\omega_3)G_{i_3}^*(\omega_4)+G_{i_3}(\omega_3)G_{i_1}^*(\omega_4)]+[G_{i_3}(\omega_1)G_{i_4}(\omega_2)+G_{i_4}(\omega_1)G_{i_3}(\omega_2)]$$

$$\times[G_{i_1}(\omega_3)G_{i_2}^*(\omega_4)+G_{i_2}(\omega_3)G_{i_1}^*(\omega_4)]$$

$$-([G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)+G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_3}(\omega_2)G_{i_4}^*(\omega_2)+G_{i_4}(\omega_3)G_{i_3}^*(\omega_2)]+[G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)]$$

$$\times[G_{i_2}(\omega_2)G_{i_4}^*(\omega_2)+G_{i_4}(\omega_2)G_{i_2}^*(\omega_2)]+[G_{i_1}(\omega_1)G_{i_4}^*(\omega_1)+G_{i_4}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)]$$

$$+[G_{i_2}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)][G_{i_1}(\omega_2)G_{i_4}^*(\omega_2)+G_{i_4}(\omega_2)G_{i_1}^*(\omega_2)]+[G_{i_2}(\omega_1)G_{i_4}^*(\omega_1)+G_{i_4}(\omega_1)G_{i_2}^*(\omega_1)]$$

$$\times[G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_1}^*(\omega_2)]+[G_{i_3}(\omega_1)G_{i_4}^*(\omega_1)+G_{i_4}(\omega_1)G_{i_3}^*(\omega_1)][G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)])$$

$$-([G_{i_1}(\omega_1)G_{i_2}^*(\omega_1)+G_{i_2}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_3}(\omega_3)G_{i_4}^*(\omega_3)+G_{i_4}(\omega_3)G_{i_3}^*(\omega_3)]$$

$$+[G_{i_1}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_1}^*(\omega_1)][G_{i_2}(\omega_3)G_{i_4}^*(\omega_3)+G_{i_4}(\omega_3)G_{i_2}^*(\omega_3)]+[G_{i_1}(\omega_1)G_{i_4}^*(\omega_1)+G_{i_4}(\omega_1)G_{i_1}^*(\omega_1)]$$

$$\times[G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)]+[G_{i_2}(\omega_1)G_{i_3}^*(\omega_1)+G_{i_3}(\omega_1)G_{i_2}^*(\omega_1)][G_{i_1}(\omega_3)G_{i_4}^*(\omega_3)+G_{i_4}(\omega_3)G_{i_1}^*(\omega_3)]$$

$$+[G_{i_2}(\omega_1)G_{i_4}^*(\omega_1)+G_{i_4}(\omega_1)G_{i_2}^*(\omega_1)][G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_1}^*(\omega_3)]+[G_{i_3}(\omega_1)G_{i_4}^*(\omega_1)+G_{i_4}(\omega_1)G_{i_3}^*(\omega_1)]$$

$$\times[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)])$$

$$-([G_{i_1}(\omega_2)G_{i_2}^*(\omega_2)+G_{i_2}(\omega_2)G_{i_1}^*(\omega_2)][G_{i_3}(\omega_3)G_{i_4}^*(\omega_3)+G_{i_4}(\omega_3)G_{i_3}^*(\omega_3)]$$

$$+[G_{i_1}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_1}^*(\omega_2)][G_{i_2}(\omega_3)G_{i_4}^*(\omega_3)+G_{i_4}(\omega_3)G_{i_2}^*(\omega_3)]+[G_{i_1}(\omega_2)G_{i_4}^*(\omega_2)+G_{i_4}(\omega_2)G_{i_1}^*(\omega_2)]$$

$$\times[G_{i_2}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_2}^*(\omega_3)]+[G_{i_2}(\omega_2)G_{i_3}^*(\omega_2)+G_{i_3}(\omega_2)G_{i_2}^*(\omega_2)][G_{i_1}(\omega_3)G_{i_4}^*(\omega_3)+G_{i_4}(\omega_3)G_{i_1}^*(\omega_3)]$$

$$+[G_{i_2}(\omega_2)G_{i_4}^*(\omega_2)+G_{i_4}(\omega_2)G_{i_2}^*(\omega_2)][G_{i_1}(\omega_3)G_{i_3}^*(\omega_3)+G_{i_3}(\omega_3)G_{i_1}^*(\omega_3)]+[G_{i_3}(\omega_2)G_{i_4}^*(\omega_2)+G_{i_4}(\omega_2)G_{i_3}^*(\omega_2)]$$

$$\times[G_{i_1}(\omega_3)G_{i_2}^*(\omega_3)+G_{i_2}(\omega_3)G_{i_1}^*(\omega_3)])\},\tag{A1}$$

where $\omega_4 = \omega_1 + \omega_2 + \omega_3$. The above equation should be interpreted as follows: if $N_t > 0$ the single summation is included, if $N_t > 1$ the double summations are included, if $N_t > 2$ the triple summation is included, and if $N_t > 3$ the quadruple summation is included.

[1] D. Brillinger, ''An introduction to polyspectra,'' Ann. Math. Stat. **36**, 1351–1374 (1965).

[2] D. Brillinger and M. Rosenblatt, ''Asymptotic theory of Kth-order spectra,'' in *Spectral Analysis of Time Series*, edited by B. Harris (Wiley, New York, 1967), pp. 153–188.

[3] M. Hinich, ''Testing for Gaussianity and linearity of a stationary time series,'' J. Time Ser. Anal. **3**, 169–176 (1982).

[4] J. Flanagan, ''Beamwidth and usable bandwidth of delay-steered microphone arrays,'' AT&T Tech. J. **64**, 983–995 (1985).

[5] C. Sydow, ''Broadband Beamforming for a microphone array,'' J. Acoust. Soc. Am. **1**, 845–849 (1994).

[6] Y. Lee, *Statistical Theory of Communication* (Wiley, New York, 1960), p. 39.

[7] J. Mendel, ''Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications,'' Proc. IEEE **79**, 278–305 (1991).

[8] R. Dwyer, ''Fourth-order spectra of Gaussian amplitude-modulated sinusiods,'' J. Acoust. Soc. Am. **90**, 918–926 (1991).

[9] D. Brillenger, ''The identification of a particular nonlinear time series system,'' Biometrika **64**, 509–515 (1977).

[10] T. Matsuoka and T. Ulrych, ''Phase estimation using the bispectrum,'' Proc. IEEE **72**, 1403–1411 (1984).

[11] K. Lii and M. Rosenblatt, ''Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes,'' Ann. Stat. **10**, 1195–1208 (1982).

[12] K. Lii and M. Rosenblatt, ''A fourth order deconvolution technique for non-Gaussian linear processes,'' in *Multivariate Analysis*, edited by P. Krishnaiah (Elsevier, Amsterdam, 1985), pp. 395–410.

[13] J. Molle and M. Hinich, ''The trispectrum,'' Proceedings of the Workshop on Higher-Order Spectral Analysis, Vail, CO, 1989 (unpublished), pp. 68–72.

[14] V. Chandran and S. Elgar, ''Statistics of tricoherence,'' IEEE Trans. Signal Process **42**, 3430–3440 (1994).

[15] R. Dwyer, ''The fourth-order cumulant spectrum of a filtered linear frequency modulated source in motion,'' J. Acoust. Soc. Am. **96**, 1460–1465 (1993).

[16] C. Nikias and M. Raghuveer, ''Bispectrum estimation: A digital signal processing framework,'' Proc. IEEE **7**, 869–891 (1987).

[17] V. Chandran and S. Elgar, ''A general procedure for the derivation of principal domains of higher-order spectra,'' IEEE Trans. Signal Process. **42**, 229–233 (1994).

[18] L. Pflug, G. Ioup, J. Ioup, and R. Field, ''Properties of higher-order correlations and spectra for bandlimited, deterministic transients,'' J. Acoust. Soc. Am. **91**, 975–988 (1992).

[19] E. Walton and I. Jouny, ''Bispectrum of radar signatures and application to target classification,'' Radio Sci. **25**, 101–113 (1990).

[20] P. Morse and K. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), Chap. 11.

[21] B. Steinberg, *Principles of Aperature and Array System Design* (Wiley, New York, 1967), pp. 71–75.

[22] R. Dwyer, ''Range and doppler information from fourth-order spectra,'' IEEE J. Ocean Eng. **16** (1991).

[23] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1989), pp. 789–792.

[24] W. Rihaczek, *Principles of High-Resolution Radar* (McGraw-Hill, New York, 1969), Chap. 1.

[25] J. Klauder, A. Price, S. Darlington, and W. Albersheim, ''The theory and design of chirp radars,'' Bell Syst. Tech. J. **34**, 745–808 (1960).

[26] R. Altes, ''Radar/sonar acceleration estimation with linear-period modulated waveforms,'' IEEE Trans. Aerosp. Electron. Syst. **26**, 914–924 (1990).

[27] D. Brill and G. Gaunaurd, ''Resonance theory of elastic waves ultrasonically scattered from an elastic sphere,'' J. Acoust. Soc. Am. **81**, 1–21 (1987).

[28] M. Junger and Feit, *Sound Structures and Their Interaction* (MIT, Cambridge, MA, 1972), pp. 318–322.

[29] W. Au, *The Sonar of Dolphins* (Springer-Verlag, Berlin, 1993), pp. 140–145.

[30] P. Woodward, *Probability and Information Theory, with Applications to Radar* (Pergamon, New York, 1953), p. 34.

[31] R. Altes, ''Some invariance properties of the wide-band ambiguity function,'' J. Acoust. Soc. Am. **53**, 1154–1160 (1973).

Efferently mediated changes in the quadratic distortion product ($f2-f1$)

Kay W. Chang[a)] and Susan J. Norton

Virginia Merrill Bloedel Hearing Research Center, Department of Otolaryngology,
University of Washington School of Medicine, Seattle, Washington 98195

In this study, a preparation was developed which continuously monitored the $f2-f1$, $2f1-f2$, and $3f1-2f2$ distortion products in guinea pigs, in response to various contralateral and ipsilateral tones and broadband noise. A consistent decline over time in the quadratic distortion product, $f2-f1$, with continuous ipsilateral stimulation was observed, confirming previous findings by Kirk and Johnstone [Hearing Res. **67**, 20–34 (1993)]. Broadband noise applied to the contralateral ear simultaneously increased both the magnitude and the rate of decline. This effect was shown for primaries across all frequencies tested ($f2$ from 2–16 kHz, $f2/f1$ ratio of 1.26), though a maximal effect was found for $f2$ of 4 kHz. Experiments using contralateral tones demonstrated a half-octave shift in the frequency specificity of this effect. A more ''immediate'' suppression effect on $2f1-f2$ (0.6 dB) was found in animals awakening from anesthesia, though this was even larger for $f2-f1$ (5.1 dB). Extended ipsilateral stimulation with contralateral noise results in completely diminishing $f2-f1$ to the noise floor. These findings suggest an adaptive process within the cochlea that may be modulated by efferent stimulation. A model relating the various even- and odd-order distortion products to outer hair cell receptor potentials is presented. © *1997 Acoustical Society of America.* [S0001-4966(97)05808-6]

PACS numbers: 43.64.Jb, 43.64.Bt, 43.64.Nf [RDF]

INTRODUCTION

In the past 15 years, research on outer hair cells (OHCs) and otoacoustic emissions (OAEs) has greatly enhanced our understanding of cochlear physiology. This work has suggested that a nonlinear active process in the cochlea, originating from motile OHCs, is responsible for the exquisite frequency selectivity and sensitivity observed in the ear (Davis, 1983; Kim, 1986; Dallos, 1992). One of the anatomical features of the cochlea that directs attention towards the OHC as the origin of active processes is its predominantly efferent innervation, as opposed to the predominantly afferent innervation to the inner hair cells (IHCs). The efferent system was first anatomically identified by Rasmussen in 1946. Since then, both the anatomy and neurochemistry of this efferent system have been characterized in a variety of animals (for review see Warr *et al.*, 1986; Altschuler and Fex, 1986; Pujol and Lenoir, 1986; Wiederhold, 1986; Bledsoe *et al.*, 1988). The efferent system can be divided into the medial and lateral olivocochlear systems (Warr and Guinan, 1978). The lateral olivocochlear (LOC) system consists of thinner, unmyelinated axons, which project predominantly to the ipsilateral side (48:15 ratio in the cat), and terminate at the peripheral processes of afferent neurons to the IHCs. In contrast, the medial olivocochlear (MOC) system innervates the OHCs. These neurons have thicker, myelinated axons, and project predominantly to the contralateral side (26:11 ratio in the cat), terminating on the modiolar side of the bases of OHCs (Warren and Liberman, 1989a). Studies in the guinea pig have indicated a 3:2 ratio of myelinated MOC

to unmyelinated LOC fibers, and a 4:1 ratio of crossed versus uncrossed OCB fibers (Terayama *et al.*, 1969; Terayama and Yamamoto, 1971).

The function of the medial olivocochlear system, however, has not been clearly demonstrated. An early hypothesis that the efferent system is required for the maintenance of normal cochlear mechanical function was shown to be false. Liberman demonstrated this in cats by sectioning the olivocochlear bundle (OCB) at the lateral border of the floor of the fourth ventricle, which effectively de-efferented both the MOC and the LOC systems to one cochlea (Liberman, 1990). This had no effect on thresholds, tuning curves, or rate-level functions. The only difference observed was a decrease in spontaneous discharge rates of auditory nerve fibers when compared to the control group. These results were confirmed by several studies in guinea pigs (Bonfils *et al.*, 1986; Littman *et al.*, 1992; Rajan *et al.*, 1990) that demonstrated no change in N_1 audiograms, CAP tuning curves, cochlear microphonic (CM), auditory brain-stem response (ABR) tuning curves, or distortion product otoacoustic emissions (DPOAEs) after transection of the OCB. None of the physiological measures of cochlear frequency selectivity was affected, and the only difference observed was in CAP masking phenomena (Bonfils and Puel, 1987).

In contrast, studies of the effects of electrically stimulating the efferent system have shown a general inhibitory effect on various physiological measures of cochlear function. In 1956, Galambos first demonstrated that electrical stimulation of the OCB could reduce the magnitude of the compound action potential (CAP) of the auditory nerve. Later studies of direct electrical stimulation have confirmed a general suppression effect on a variety of cochlear and eighth

---
[a)]Electronic mail: kchang@u.washington.edu

nerve potentials (Fex, 1962; Desmedt, 1962; Dewson, 1967; Wiederhold and Kiang, 1970; Nieder and Nieder, 1970a, b; Konishi and Slepian, 1971; Teas *et al.*, 1972; Brown *et al.*, 1983; Brown and Nuttall, 1984; Guinan and Gifford, 1988a, b).

Since the MOC system innervating the OHC is predominantly contralateral, contralateral acoustical stimulation was considered as another possible method of specifically stimulating the efferent system. A number of studies demonstrated suppression of various ipsilateral measures of auditory function in the presence of a contralateral acoustic stimulus (Buño, 1978; Murata *et al.*, 1980; Folsom and Owsley, 1987; Liberman, 1989; Warren and Liberman, 1989a, b). This suppression effect was completely eliminated with transection of the MOC (Liberman, 1989; Warren and Liberman, 1989a, b). Using contralateral tones, Warren and Liberman observed a maximal effect when the contralateral tone was near the characteristic frequency (CF) of the measured auditory nerve fiber, which decreased as the suppressor was moved away from the CF, indicating that the efferently mediated suppression had some frequency selectivity.

Although their exact origin is unknown, a large body of data indicates that otoacoustic emission generation, at least in mammals, is dependent upon OHC integrity (Kemp, 1978; Davis, 1983; Kim, 1986; Dallos, 1992). Distortion product otoacoustic emissions (DPOAEs), where two tones at frequencies $f1$ and $f2$ ($f1 < f2$) are introduced into the ear canal, have proven particularly useful in exploring OHC function in a frequency-specific manner. By monitoring DPOAEs during electrical stimulation of the OCB, Mountain (1980) demonstrated suppression of the quadratic difference tone (QDT), $f2 - f1$. Changes in both $f2 - f1$ and $2f1 - f2$, the cubic distortion tone (CDT), were demonstrated by Siegel and Kim (1982); also using direct electrical stimulation of the OCB. Contralateral suppression of several classes of OAEs have been demonstrated in humans. These include spontaneous otoacoustic emissions (SOAEs) (Grose, 1983; Schloth and Zwicker, 1983; Rabinowitz and Widin, 1984; Mott *et al.*, 1989; Kujawa and Glattke, 1989, 1990; Moulin *et al.*, 1992; Harrison and Burns, 1993), transiently evoked, or click-evoked otoacoustic emissions (TEOAEs) (Collet *et al.*, 1990, 1992; Ryan *et al.*, 1991; Veuillet *et al.*, 1991, 1992; Berlin *et al.*, 1993b; Rossi *et al.*, 1993; Williams *et al.*, 1993, 1994), as well as DPOAEs (Brown and Norton, 1990; Moulin *et al.*, 1992, 1993). In animals, DPOAEs have been shown to undergo a similar suppression with contralateral acoustic stimuli. Guinea pigs subjected to contralateral white noise were observed to have a reduction in the CDT that was not observed in animals that had their MOC cut by a midline sagittal section at the floor of the fourth ventricle (Puel and Rebillard, 1990). Kujawa *et al.* (1993) provided evidence for a cholinergic efferent mechanism when they used cochlear perfusion of strychnine, curare, and atropine to reversibly block the suppression of the CDT from contralateral sound stimulation. Contralateral acoustic suppression of DPOAEs have also been demonstrated in cats (Warren and Liberman, 1989a, b) and rabbits (Whitehead *et al.*, 1991b), although the latter attributed the effect to the acoustic reflex.

One additional phenomenon, which has recently proven to be a promising line of investigation into efferently mediated effects, is the modulation of the QDT, $f2 - f1$, by various ipsilateral and contralateral stimuli. This was first observed by Brown (1988) in guinea pigs and gerbils. With continuous low-level stimulation using two primary tones $f1$ and $f2$, large reductions of up to 15 dB below control levels were observed in the amplitude of emitted $f2 - f1$. Recovery to pre-exposure control levels occurred when the stimulation ceased. The magnitude of $f2 - f1$ reduction observed was thought to be affected by the level of anesthesia, and recovery was also delayed when the contralateral ear was stimulated with tones during the recovery period, which led to the hypothesis that this effect was a result of efferently mediated modulation of cochlear mechanics. Whitehead *et al.* (1991a) observed a similar slow reversible decline of $f2 - f1$ during continuous primary stimulation in unanesthetized rabbits. Not observing a delay in recovery with stimulation of the contralateral ear, they did not attribute this phenomenon to efferent influences. Both of these studies indicated that the CDT, $2f1 - f2$, remained relatively stable in contrast to the large reductions seen in the QDT. The stability of the CDT as well as the low stimulation levels (below 55 dB) made acoustic fatigue an unlikely explanation for the effect.

Kirk and Johnstone (1993) studied these suppressive effects on $f2 - f1$ extensively in guinea pigs, demonstrating separate ipsilateral and contralateral effects which were additive. In their experiments, continuous stimulation of the test ear with primary tones, or with single tones near the primary frequencies, caused a reduction in amplitude of $f2 - f1$ but not $2f1 - f2$, similar to the two previous studies. Broadband noise applied to the contralateral ear in addition to the continuous primary stimulation increased the suppression obtained in $f2 - f1$. The magnitude of the effect induced by the ipsilateral stimulation alone added to the magnitude of the effect induced by the contralateral stimulation alone were roughly equivalent to the effect of combined ipsilateral and contralateral stimulation. Systemic administration of strychnine was shown to block the contralateral suppression effect, but the ipsilateral suppression remained. Intracochlear perfusion of bicuculline, a GABA-specific antagonist, blocked both contralateral and ipsilateral suppression. Kirk and Johnstone further observed that these effects could only be found in the frequency range between 2 and 7 kHz. They concluded that this phenomenon was mediated by a GABA-ergic efferent system in the apical cochlea.

In the present study, an experimental preparation was developed to allow continuous monitoring of $f2 - f1$, $2f1 - f2$, and $3f1 - 2f2$ in an extremely reliable and accurate fashion, while systematically studying the previously mentioned contralateral and ipsilateral suppression phenomenon. Guinea pigs were used, and Kirk and Johnstone's basic findings were confirmed, and extended. The phenomenon was found to be present across a broader range of frequencies (from 2 to 16 kHz). Additional frequency-specific information was also obtained, confirming a tuned effect which incorporated a half-octave shift. One final question that was resolved in these experiments was why Kirk and Johnstone were unable to demonstrate a suppressive effect on $2f1$

$-f2$ using contralateral acoustic stimulation, a finding which has been observed in the same animal model in two previous studies, as well as numerous studies in other animals and in humans.

I. METHODS

A. Animal preparation

Adult Hartley strain albino guinea pigs weighing 300–400 g were obtained from a commercial animal supplier (Simonsen Laboratories, Gilroy, CA). Animal preparation and recording was performed in an I.A.C. double-walled acoustic booth. An intraperitoneal injection of a mixture of ketamine hydrochloride (Ketaset 80 mg/kg) and xylazine (Rompun 10 mg/kg) was used to induce anesthesia. Since the average experimental session lasted three hours, additional injections of ketamine (40 mg/kg) and xylazine (5 mg/kg) were given as necessary to maintain anesthesia (usually only one additional injection was required). Subjects were placed in a custom head holder and a thermocouple was inserted into the rectum to monitor core temperature. A microprocessor-controlled heating pad underneath the animal maintained its temperature at 35 °C–37 °C. Under microscopic visualization, a probe consisting of a microphone assembly (Etymotics ER-10B) and two transducer tubes was coupled snugly to one ear canal. An identical probe was also coupled to the opposite ear canal.

Procedures for the care and use of the animals reported on in this study were approved by the University of Washington Animal Care Committee.

B. System calibration, stimulus generation, and emission recording

The system for generating the two probe tones and detecting emissions was the same as that described in previous experiments performed in our laboratory (Mills *et al.*, 1993; Mills and Rubel, 1994; Chang and Norton, 1996). Briefly, it consisted of a host computer (Macintosh IIfx), a digital signal processing (DSP) board (Spectral Innovations), and three channels for analog-to-digital conversion and two for digital-to-analog (DA) conversion. Two tones generated by the DA converters (frequencies $f1$ and $f2$) were sent through a remote-controlled attenuator to two Beyers DT-48 transducers. The resulting stimuli were delivered to the ear canal through two small tubes that terminated in a molded probe tip at the end of the Etymotics ER-10B microphone assembly. The entire assembly was positioned with a micromanipulator, which allowed the probe tip to be inserted tightly into the ear canal and maintained in a fixed orientation. With this system, stimulus signal levels of at least 85 dB SPL could be obtained up to 16 kHz. Closed cavity tests showed detectable instrumental distortion was at least 75 dB below the stimulus levels.

At the beginning of each recording session, a probe tube microphone (Etymotics ER-7C) was connected temporarily to one of the small tubes leading through the ER-10B microphone assembly, in place of one of the transducers. This tube was the one for which the ER-7C had been calibrated by the manufacturer. A wide-band noise signal was delivered to the ear canal through the other small tube by the remaining transducer and both microphone outputs were fast Fourier transformed (FFT). A microphone correction table was generated on the computer to correct the amplitude output of the ER-10B according to the output of the ER-7C. The ER-10B microphone output was then corrected in real time after each FFT by use of this table. This gave an absolute amplitude measurement estimated to be correct to ± 1 dB, as measured at the end of the probe tube.

For every measurement, averaging for the FFT was begun after a delay of 50 ms, to avoid startup transients. The microphone output was first averaged for 100 ms. If the measured amplitudes of the primaries deviated more than 1 dB from the programmed values, the attenuator settings were corrected. The microphone signal was then averaged for a programmed length of time ranging from 2 to 10 s, as determined by a parameter list loaded into the system. Stimulus parameters used in this study were $f2/f1 = 1.26$, $L1$ (amplitude of $f1$)$=65$ dB SPL and $L2$ (amplitude of $f2$)$=55$ dB SPL. Amplitude and phase of the FFT component at the two stimulus frequencies and at the following frequencies were recorded for each stimulus presentation: $2f1-f2$, the third-order term usually known as the cubic distortion tone (CDT); $3f1-2f2$, the fifth-order term having, like the CDT, an emission frequency below the two primaries; $2f2-f1$, the third-order term with an emission frequency above both primaries; and $f2-f1$, the quadratic difference tone (QDT). The noise floor was calculated by averaging the amplitudes obtained from the spectrum analysis in a 50-Hz window bracketing the distortion tone for both $2f1-f2$ and $f2-f1$.

C. Contralateral stimulus generation and measurement

Contralateral stimuli were generated by a Hewlett–Packard 8904A Multifunction Synthesizer, amplified through a RD Systems amplifier, and applied to a Beyers DT-48 transducer, which was coupled to the ear via one of the transducer tubes in the contralateral probe assembly. This assembly was identical to the ipsilateral probe and positioned in likewise fashion to the opposite ear, with the probe tip coupled tightly to the ear canal and maintained in a fixed orientation by the micro-manipulator. Recordings from the Etymotics ER-10B microphone in the probe were analyzed by a Hewlett–Packard HP3561A dynamic signal analyzer. The contralateral noise stimuli used in this experiment were broad band with an average overall level of 70 dB SPL. Spectrum analysis of the noise indicated that it effectively extended from 250 to 7500 Hz within ± 5 dB SPL level with low- and high-frequency roll-offs of approximately 12 and 17 dB per octave, respectively. Contralateral tone stimuli were also generated by the multifunction synthesizer. The amplitudes of these stimuli were 70 dB SPL, with the exception of the 16-kHz tone, which was set to 80 dB SPL.

D. Experiment

The data presented in this study were accumulated from 53 separate experimental sessions run on 16 different guinea
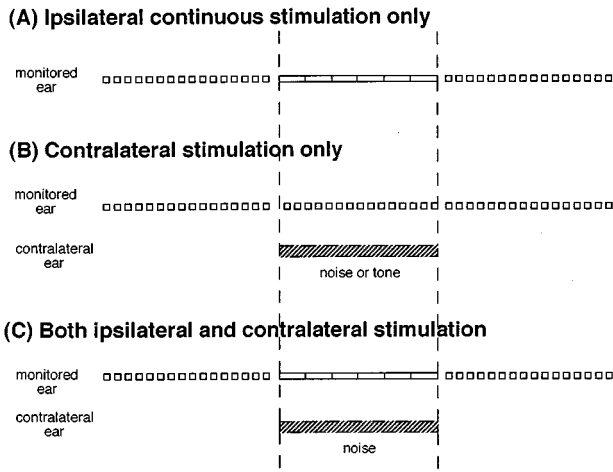
FIG. 1. Diagram of the three stimulation paradigms. The small boxes spaced apart from each other represent presentation of brief (2 s) intermittent stimuli. The long rectangles connected to one another represent presentation of the continuous stimuli. The horizontal cross-hatched bar represents the contralateral stimulus.

FIG. 2. The DPOAE recordings across time for a guinea pig receiving continuous ipsilateral stimulation with $f1 = 3.17$ kHz and $f2 = 4$ kHz.

pigs. Baseline DPOAEs were obtained from each guinea pig across a broad range of frequencies to ensure that they fell within normal baselines as established by previous experiments conducted at our laboratory. The DPOAEs were monitored across time by presenting 2-s-long stimulus pairs intermittently, while averaging the recordings from the microphone during stimulus presentation. Eight stimulus pairs with $f2$ ranging from 1.41 to 16 kHz in half-octave steps were presented with at least 2 s pause between each presentation and from 18 to 32 s pause between each set of eight stimulus presentations. This allowed a broad range of frequencies to be monitored with a stimulus exposure schedule which avoided presentation of the same frequencies for more than 4 s each minute. In Fig. 1, the intermittent stimuli are represented by the small boxes. Continuous ipsilateral sound stimulation was obtained by presenting 10-s-long stimulus pairs continuously, with the resulting DPOAE recordings averaged across this time. Since the DSP board used to generate the two stimuli and to record the DPOAEs required a brief pause in stimulus presentation in order to also perform analysis of the emission data, continuous stimulation was closely approximated by adapting system parameters so that only a minimal pause time occurred between each 10-s-long stimulus pair. This averaged to be approximately 62 ms. Thus, a 10-min long ''continuous'' presentation could be closely approximated by 60 10-s-long stimulus pairs strung end to end (which actually results in 10 min 3 s of elapsed time). In Fig. 1, continuous stimuli are represented by the long rectangles joined together end to end.

The top of Fig. 1 illustrates the ipsilateral-continuous-stimulation-only experiments. Intermittent stimuli are given to monitor baseline DPOAEs, then a period of continuous stimuli (usually 10 min long), with recovery monitored by intermittent stimuli again. The middle of Fig. 1 illustrates the contralateral-stimulation-only experiments. In these experiments, intermittent stimuli are given throughout. After a baseline period, the contralateral noise or tone is turned on for 10 mins, after which recovery is monitored. The bottom
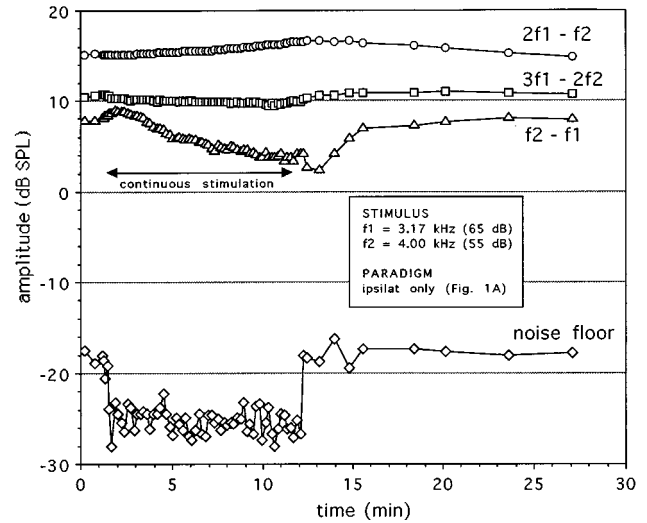
of Fig. 1 illustrates the combined continuous stimulation with contralateral noise experiments. These experiments were run identically to the ipsilateral-only experiments with the addition of the contralateral noise during the period of the continuous ipsilateral stimulation.

II. RESULTS

A. Individual data

The DPOAE data obtained from the preparation described were extremely stable and reproducible throughout each and every recording session for every animal. The DPOAE amplitudes were recorded by the system to the nearest 0.1 dB. Initial pilot experiments monitoring the various distortion products intermittently without any continuous or contralateral sound stimulation showed stable measurements across hour-long recordings, with variations of no more than 0.3 dB.

Because the results of every experiment were so similar, raw unprocessed data from a few typical examples to illustrate some points will be discussed first, before presenting a statistical summary of the data. Figure 2 illustrates the raw unprocessed data from one guinea pig subjected to a 10-min-long continuous stimulation of a 3174-Hz tone ($f1$) at 65 dB and a 3998-Hz tone ($f2$) at 55 dB. Amplitudes of DPOAEs at $f2 - f1$, $2f1 - f2$, and $3f1 - 2f2$ are shown, along with the noise floor calculated around $f2 - f1$. Noise floor calculations performed at $2f1 - f2$ were invariably lower, since $f2 - f1$ was always at a lower frequency than $2f1 - f2$, which placed it in a region with higher noise. Thus, in this and all following graphs, the noise floor data will be taken from the noise floor calculated around $f2 - f1$. During continuous stimulation of the primary tones, an essentially continuous plot of DPOAEs averaged over 10-s intervals was obtained over the entire 10-min period. The DPOAEs monitored before and after the continuous stimulation were averaged over intermittent 2-s stimulus presentations, as described in Sec. I. Note that the noise floor is about 7 dB lower during the continuous stimulation. This is the result of
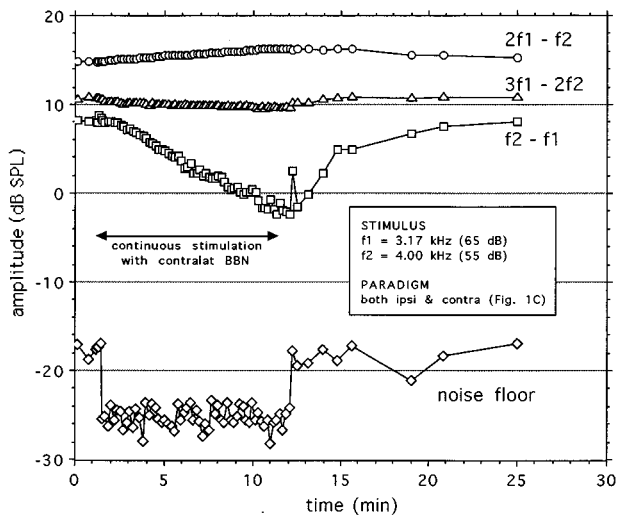
FIG. 3. The DPOAE recordings across time for the same guinea pig shown in Fig. 2. A contralateral noise stimulus is being presented simultaneously with presentation of the continuous ipsilateral 3.17 and 4 kHz primaries.

FIG. 4. The DPOAE recordings across time for a guinea pig receiving continuous ipsilateral stimulation with 1.59 and 2 kHz primaries with simultaneous contralateral noise.

the linear time-averaging algorithm used to calculate the FFT, in which a doubling of the integration time lowers the noise floor by 3 dB (in this particular case, the integration times during continuous stimulation were five times longer, and thus $10 \log 5 \approx 7$ dB, which was the decrease observed in the data). During the continuous primary presentation, $2f1 - f2$ and $3f1 - 2f2$ remain relatively stable, though with a hint of a slight increase in $2f1 - f2$ and a slight decrease in $3f1 - 2f2$. In contrast, $f2 - f1$ shows a 3-dB decrease in amplitude over the first 5 min of continuous stimulation, after which the trajectory of the decline flattens somewhat to a 1-dB decrease over the second 5 min for a total 4-dB decrease in amplitude over 10 min of continuous stimulation. After the continuous stimulation ceased, intermittent DPOAE measurements showed a recovery of $f2 - f1$ back to baseline on a timescale of about 4 min. This pattern of behavior (i.e., rapid initial decline with leveling off, and complete recovery after cessation of stimulation) was closely reproduced in every ipsilateral experiment that was performed (41 trials).

In Fig. 3 are data from the same guinea pig as in Fig. 2. The same 3174/3998-Hz stimuli were given continuously for 10 min, with the addition of a 70 dB SPL broadband noise (BBN) applied to the contralateral ear during the 10-min-long continuous ipsilateral stimulation. $2f1 - f2$ and $3f1 - 2f2$ behave in similar manner as before. $f2 - f1$ shows a decline in amplitude, with the trajectory of this decline remaining constant for the entire ten minutes, resulting in a total decline of almost 10 dB. After cessation of both the contralateral noise and the continuous primaries, a recovery of $f2 - f1$ back to baseline was observed on a timescale of about 6–8 min. This pattern of behavior (i.e., a relatively constant and steeper decline, with recovery after cessation of stimulation) was observed in every combined ipsi contralateral experiment (57 trial runs). An important point to note in this graph is that the noise floor is at the same level (approximately -25 dB) as the previously presented run which did not have a contralateral noise stimulus. Spectrum analysis in
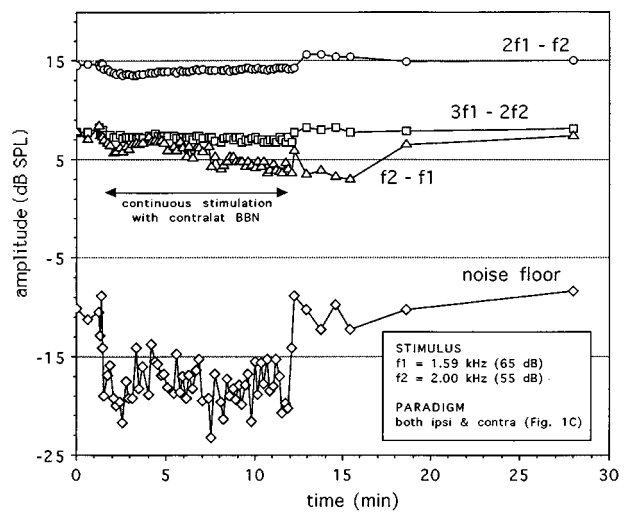
all runs, examined carefully for cross-over effects of contralateral noise, failed to show any change in noise floor with contralateral noise, or any evidence of signal crossing over above the noise floor of the measurement system, when investigating contralateral tones.

The suppression effect of ipsilateral or combined ipsi/contralateral stimulation was a robust finding, and was obtained in 100% of all attempted experiments with stimulus pairs across the entire frequency spectrum. Figure 4 represents one ear tested at the low end of the frequency range examined, with $f1 = 1587$ and $f2 = 1996$ Hz. A similar suppression in $f2 - f1$ is observed, though the total effect was weaker than that observed with the 3174/3998 stimuli. Note that the noise floor of this run is considerably higher (-17 vs -25 dB) than the previously two presented runs. $f2 - f1$ in this case was at 409 Hz, where the noise floor of the emission recording system began to increase sharply. In fact, a steep increase in noise was observed for frequencies below 400 Hz, which precluded experiments using lower frequency stimuli. The noise floor observed in Fig. 5, on the other hand, is observed to be considerably lower (-30 vs -25 dB) than that observed in Figs. 2 and 3. This was a run from an ear tested at the high end of the frequency range, with $f1 = 12\,695$ and $f2 = 15\,997$ Hz.

B. Averaged data

Fifty-three trial runs were performed using the ipsilateral and combined ipsi/contralateral stimulation paradigms. The figures presented are typical of all of these runs. Table I summarizes the mean suppression effects obtained for each of the stimulus pairs tested. Standard deviations are provided for mean data accumulated from three or more test trials. Included on this table are only the standardized runs lasting 10 min, representing half of all the data obtained, which included many shorter, and a few longer runs. While a suppression effect was obtained from across all frequencies presented, it was larger in the middle frequencies, from 2.3–6.7 kHz.
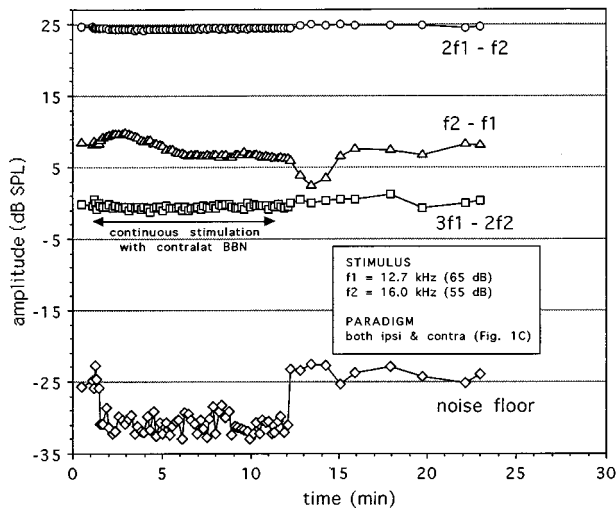
FIG. 5. The DPOAE recordings across time for a guinea pig receiving continuous ipsilateral stimulation with 12.7 and 16 kHz primaries with simultaneous contralateral noise.

FIG. 6. Averaged normalized data for animals receiving continuous ipsilateral stimulation with 3.17 and 4 kHz primaries with (b) ($n = 13$) and without (a) ($n = 12$) contralateral broadband noise (BBN). The vertical dashed lines represent the beginning and end of the continuous stimuli. Error bars represent one standard deviation from the mean.

The suppression effect from the 3174/3998-Hz stimulus pairs was studied more extensively, since it consistently demonstrated the strongest effect. Data accumulated from many individual runs were analyzed in the following fashion. The data from each animal were first truncated so that 2 mins of intermittently obtained DPOAE data remained before the ipsi or ipsi/contra stimulation and 7 min remained afterwards. This approximately time aligned runs from different animals. The data for each animal were normalized by subtracting the amplitude of the first DPOAE data point from that animal from all subsequent recordings from that animal. The data from all animals were then accumulated in 27 time bins and averaged. The processed data from 12 animals given continuous ipsilateral stimulation and 13 animals given continuous ipsilateral stimulation with simultaneous contralateral noise are presented in Fig. 6. For purposes of clarity, error bars (representing one standard deviation from the mean) are drawn only on the CDT ($2f1 - f2$) and QDT ($f2 - f1$) plots (standard deviations for the $3f1 - 2f2$ data were similar to those for the CDT data). Across animals, a slight increase of about 0.5 dB in the CDT was observed. The accumulated processed $f2 - f1$ data are essentially identical to the raw $f2 - f1$ data presented in Figs. 2 and 3. The suppression effect was extremely characteristic and repro-

ducible across all animals, with standard deviations of only 1 dB after ten minutes of stimulation.

As mentioned in Sec. I, during intermittent monitoring of the DPOAE before and after stimulation, distortion products from other frequency pairs could also be monitored. In Fig. 7, $f2 - f1$ (for $f2$ sweeping from 2–16 kHz in half-octave steps) is plotted in decibels below control for the two stimulation paradigms, immediately after cessation of the stimulation. Error bars again represent one standard deviation from

TABLE I. Mean decline in $f2 - f1$ emission after ten minutes of continuous stimulation.

$f1$ (kHz)	$f2$ (kHz)	Ipsilat stim only		Ipsilat w/ contra BBN	
		# trials	Mean decline (dB±s.d.)	# trials	Mean decline (dB±s.d.)
1.59	2	2	2.8	3	3.3 (± 1.0)
1.89	2.37	0		1	5.3
2.24	2.83	1	3.8	2	6.1
3.17	4	12	4.3 (± 1.1)	13	8.1 (± 1.3)
4.49	5.65	0		3	6.9 (± 1.2)
5.33	6.73	1	3.5	1	6.4
6.35	8	3	2.4 (± 0.8)	2	1.8
8.98	11.3	0		2	2.1
12.7	16	2	1.1	3	2.5 (± 0.6)
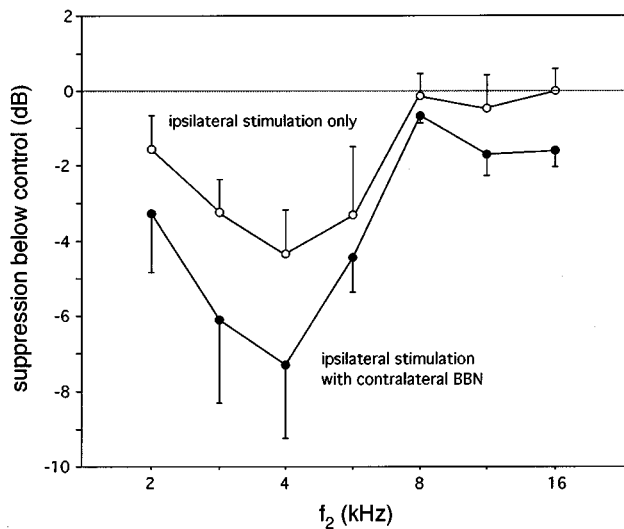
FIG. 7. The amount of suppression in $f2-f1$ (dB) referenced across the $f2$ primary after cessation of a 10-min-long ipsilateral ($n=12$) and combined ipsi/contralateral ($n=13$) stimulation.

the mean. A similar tuned effect is seen for both stimulation paradigms, with the exception that the combined ipsi/contralateral stimulation resulted in suppression at 8 kHz and upwards while ipsilateral stimulation only did not result in suppression beyond 8 kHz.

C. Contralateral stimulation only data

The experiments involving contralateral stimulation only, while intermittently monitoring the ipsilateral side, allowed DPOAEs across the frequency spectrum to be monitored during, as well as before and after, the stimulation. Figure 8 presents the averaged normalized $f2-f1$ data from 12 animals during contralateral stimulation with a 70 dB SPL level BBN (error bars not drawn). A similar procedure was used to process the data except that fewer, wider time bins were used since the intermittently recorded points were fewer and spaced farther apart. Each plot is in decibels below control, and labeled by the $f2$ primary.

A general suppression effect is seen in $f2-f1$ across all of the $f2$ stimuli. Plotting the amplitude of the effect across frequencies at the point of maximal suppression (i.e., the last measurement point before turning off the contralateral noise, as marked by the arrow at the bottom of Fig. 8) results in the top panel of Fig. 9. There was about a 3-dB suppression effect from 1.4–4 kHz, which rolled off to a 1-dB effect at 16 kHz.

On the lower panels in Fig. 9 are plotted the maximal suppression obtained after 10 min of contralateral stimulation with various tones. All of the tones were 70 dB SPL in amplitude, except the 16-kHz tone, which was 80 dB SPL (we used 80 dB SPL at 16-kHz because the single trial run with a 70 dB SPL 16-kHz tone failed to show any suppression). The 1- and 2-kHz tones resulted in the strongest suppression, and the 16-kHz tone the weakest. A particularly intriguing result was that the contralateral stimulus results in a frequency-specific suppression of $f2-f1$, which is broadly tuned near a half-octave above $f2$. This was best seen in the
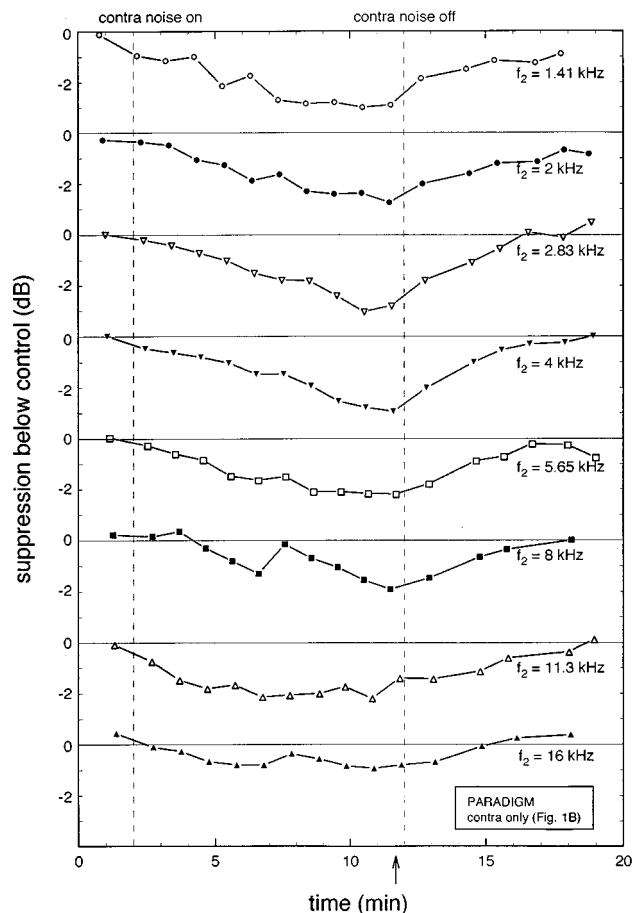
FIG. 8. Averaged normalized data ($n=12$) for animals receiving contralateral noise only. Data are obtained from intermittent stimulus pairs, and plotted in separate panels by the $f2$ stimulus in decibels below control. Each division on the y axis represents one decibel. The vertical dashed lines represent the beginning and end of the contralateral noise. The arrow at the bottom is pointing to the time of the last measurements obtained before turning off the contralateral noise.

4- and 8-kHz data, which resulted in narrower effects. Due to the broad effect seen with 2-kHz stimulation, the frequency of the point of maximum suppression is not as obvious. Superimposed on the 2-kHz panel is a dashed line, which represents the averaged data after only 2 min of the 10-min-long contralateral tone. This suggests that an approximately half-octave shift is present for the 2-kHz tone as well. The $2f1-f2$ and $3f1-2f2$ data (not shown in Fig. 9) across all frequencies were within a standard deviation of the control, or zero level.

D. Effects of extended contralateral stimulation

Examining the results of individual animals exposed to continuous ipsilateral primaries showed a consistent leveling off of the $f2-f1$ suppression, almost but not quite to an ''asymptote.'' The averaged data in Fig. 6 does not show this asymptotic phenomenon very well since the different animals demonstrated a leveling-off of their decline at varied time points between 5 and 8 min of stimulation. The averaged curve thus smoothed out the asymptote that was observed in the individual data. A 20-min-long run performed in two animals demonstrates that the suppression of $f2$

K. W. Chang and S. J. Norton: Efferent changes in DPOAEs
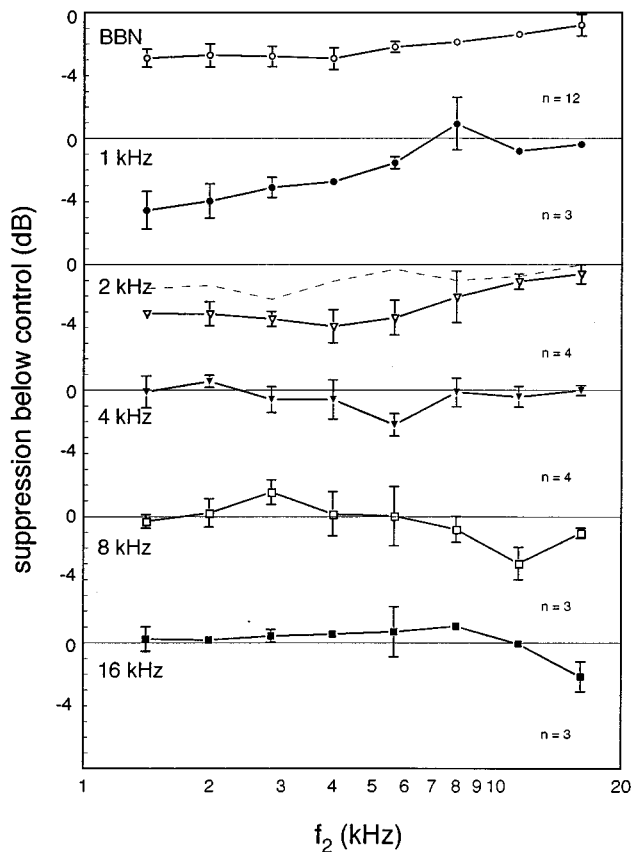
FIG. 9. Top panel depicts the total amount of suppression in $f2-f1$ obtained after 10 mins of contralateral noise, the time point marked by the arrow at the bottom of Fig. 8. The x axis plots the $f2$ stimulus primary. Each division on the y axis represents one decibel below baseline. Each succeeding panel depicts suppression induced by 10 mins of contralateral tone, the frequency being indicated in the upper left of each panel. The number of animals tested in each test condition is indicated in the lower right corner. The dashed line on the 2-kHz contralateral tone panel represents the amount of suppression after 2 mins of stimulation.

$-f1$ indeed levels off during continuous ipsilateral stimulation. Figure 10 shows the results from one animal exposed to 3174/3998-Hz continuous primaries (the results from the other animal were essentially identical).
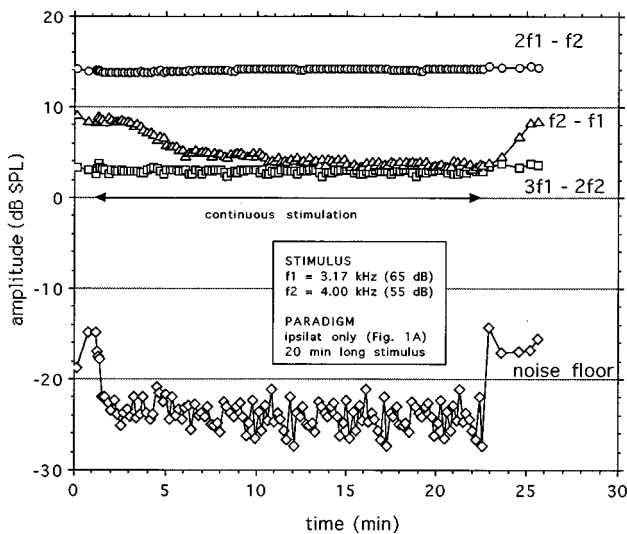
FIG. 10. The DPOAE tracing across time for a guinea pig receiving a 20-min-long continuous ipsilateral stimulation with 3.17 and 4 kHz primaries.
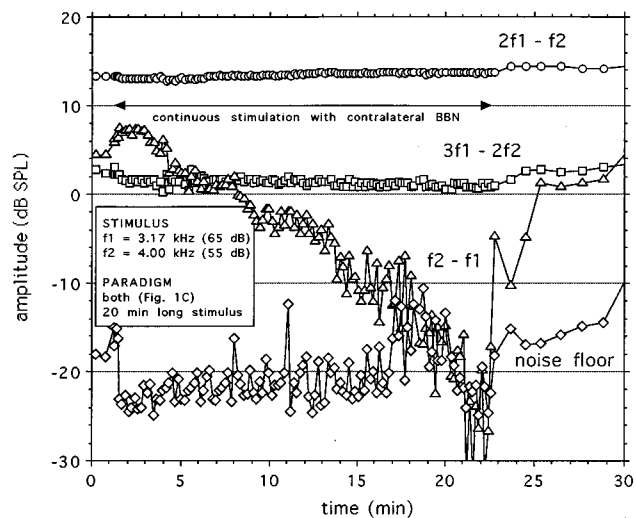
FIG. 11. The DPOAE tracing across time for a guinea pig receiving a 20-min-long contralateral noise stimulus simultaneously with presentation of continuous ipsilateral 3.17 and 4 kHz primaries.

In contrast, during extended 20-min-long stimulation of ipsilateral 3174/3998-Hz primaries including contralateral noise stimulation, the decline in $f2-f1$ continued on the same trajectory all the way down into the noise floor, as shown in Fig. 11. In this animal, $f2-f1$ was suppressed by over 30 dB, while $2f1-f2$ and $3f1-2f2$ remained relatively stable in comparison. Particularly notable, though, is that $f2-f1$ recovers back to baseline in the same 6-min timescale that was observed in the 10-min-long experiments. Thus, even though the QDT was essentially eliminated by the suppression effect, it recovered back to normal levels in the same amount of time. A second animal given combined ipsi/contralateral stimulation extended for 20 min duplicated these results, with suppression of $f2-f1$ down into the noise floor, and recovery back to baseline 7 min after cessation of stimulation.

E. Effects of anesthesia on contralateral suppression

In the experiments described above, the suppression of $2f1-f2$ by contralateral noise as demonstrated by Puel and Rebillard and by Kujawa and associates was not observed. Instead, $2f1-f2$ was shown to be relatively stable, and, if anything, slightly increased in amplitude in response to the contralateral stimuli. Further investigations revealed that suppression of $2f1-f2$ could indeed be consistently demonstrated, if the animal was emerging from anesthesia.

During initial investigations of the contralateral suppression effect on $2f1-f2$, in experiments designed similar to the ones described in the Puel and Rebillard and Kujawa *et al.* papers, a suppression of $2f1-f2$ was sporadically observed, but, at that time, the reasons were not understood. Failure to consistently show suppression of $2f1-f2$ eventually led to the experiments described in this paper, which show a robust, extremely consistent, and easily demonstrable effect. Continuing to look for $2f1-f2$ suppression, during the approximately 3-h experimental sessions which provided
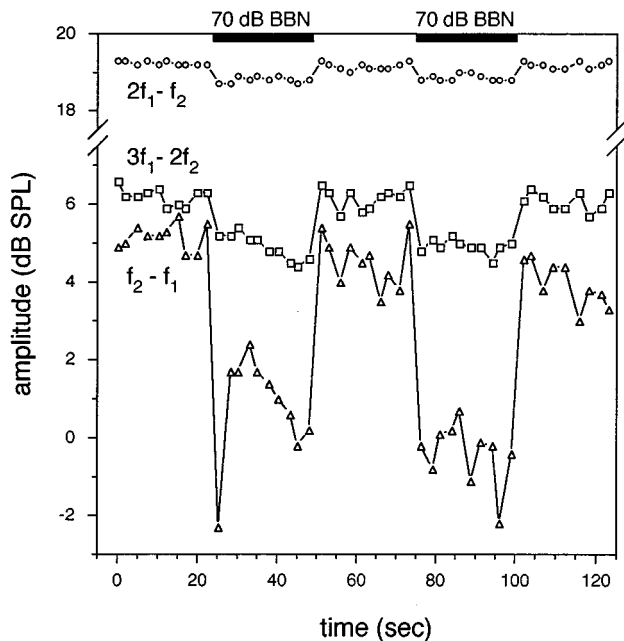
FIG. 12. Continuous DPOAE tracing across time for a ''light'' guinea pig just emerging from anesthesia. Two-second-long 3.17 and 4 kHz stimulus primaries (at 65 and 55 dB) were used to obtain the emissions. The dark bars at the top of the graph indicate periods of 70 dB SPL BBN stimulation to the contralateral ear.

FIG. 13. Continuous DPOAE tracing across time for a ''light'' guinea pig given a contralateral BBN increasing in level from 70 to 90 dB SPL. Two-second-long 3.17 and 4 kHz stimulus primaries (at 55 and 45 dB) were used to obtain the emissions. The dark bars at the top of the graph indicate periods of contralateral stimulation.

the data for this paper, animals were periodically continuously monitored for 2–3 min with 3174/3998-Hz primaries while turning a contralateral stimulus on and off, to see if the $2f1-f2$ suppression that was sporadically observed in previous experiments could be reproduced. Suppression of $2f1-f2$ was observed in every animal tested, but only at particular times, which were sometimes in the middle of the session (about the time that an additional maintenance dose of anesthesia was given) and always at the end of the experiment, when the animals were beginning to emerge from anesthesia. Data like those in Fig. 12 were consistently reproduced, though extended measurements could never be performed at this point, because the animal would require additional anesthesia.

The average suppression of $2f1-f2$ observed across animals ($n=16$) during 70-dB contralateral BBN was 0.6 dB with a standard deviation of ± 0.3 dB. Suppression of $3f1-2f2$ and particularly $f2-f1$ were of much greater magnitude—1.8 ± 0.7 and 5.1 ± 1.3 dB, respectively. The suppression effect increased with increases in intensity of the contralateral BBN, as demonstrated in Fig. 13. One feature particularly notable about this effect is that it is an immediate effect. In contrast to the slow reversible decline in $f2-f1$ occurring over a timescale of minutes, the suppression of DPOAEs observed in awake animals occurred immediately upon application of the contralateral stimulus (note that unit on the x axis of Figs. 12 and 13 is in *seconds* of elapsed time). Another interesting feature of the suppression, was that the $f2-f1$ measurement taken in the initial second of contralateral noise, often, though not always, exhibited a sharp drop lower than subsequent measurements taken in the following seconds.
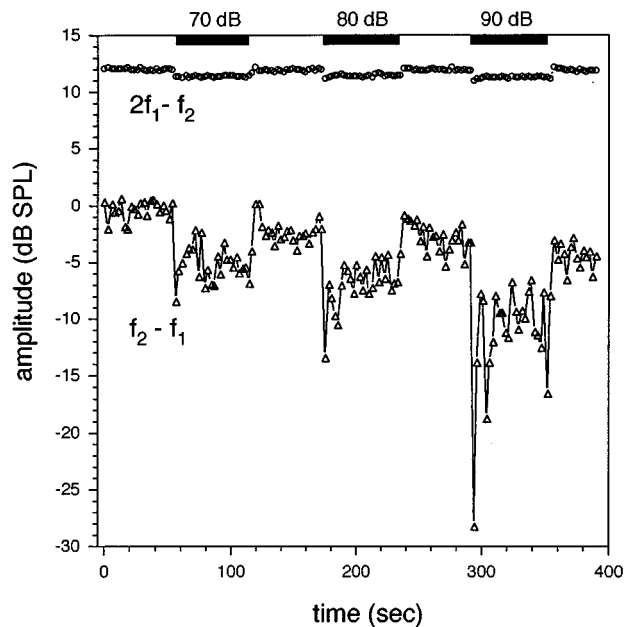
Measurements on two animals anesthetized with ketamine only (80 mg/kg) without xylazine were attempted. These indicated that the immediate contralateral suppression effect was indeed present, which suggested that xylazine was the substance which prevented observation of this effect during periods of deep anesthesia.

III. DISCUSSION

A. Comparisons with previous data

In these experiments, Kirk and Johnstone's findings on efferently mediated modulation of $f2-f1$ in response to contralateral and ipsilateral acoustic stimulation were confirmed in the guinea pig. There were a few differences between our findings, which were probably due to differences in stimulus parameters and in the sensitivity of our DPOAE recording systems.

The data obtained in the experiments described in this paper were accurate to the tenth of a decibel. Close examination of the figures presented by Kirk and Johnstone indicated all of their measurements to be rounded into units of one decibel. With contralateral and/or ipsilateral stimulation, Kirk and Johnstone observed $2f1-f2$ to be unchanged while we found a small but consistent increase in amplitude of $2f1-f2$ by about half a decibel, an amount that would have been not picked up in Kirk and Johnstone's measurements. Interestingly, a similar small increase in $2f1-f2$ by 0.4 dB was observed by Whitehead and associates (1991a) in rabbits given 30 min continuous ipsilateral stimulation. $3f1-2f2$ was observed to be relatively stable as well, though a small decrease in amplitude of half a decibel was observed in the combined ipsi/contralateral stimulation condition.

1727   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

K. W. Chang and S. J. Norton: Efferent changes in DPOAEs   1727

A decline in the QDT was induced with primaries up to 16 kHz in the experiments described here. This is in contrast to Kirk and Johnstone's work, where the effect was generated only by primary tones between 2 and 7 kHz. Several comments should be made about this difference since Kirk and Johnstone make a particular point about the apical distribution of GABA-related staining of efferent terminals on OHCs. Reviewing Table I, one can note that the strongest suppression effects observed in our experiments were in the frequency regions between 2 and 7 kHz, with a much weaker, though definitely present, effect observed at higher frequencies. The data presented by Kirk and Johnstone appear to all be rounded to the nearest decibel, and may not have picked up these smaller effects. Another possible reason for the difference between our findings and theirs may be differences in stimulus parameters. The primaries utilized by Kirk and Johnstone were equilevel, with a $f2/f1$ ratio of 1.17. In contrast, our experiments utilized $f2/f1$ frequency ratios of 1.26 and unequal level primaries such that $L2 = L1 - 10$ dB. These parameters have been found to produce robust emissions in guinea pigs (Brown, 1987; Brown and Gaskill, 1990; Brown, 1993) and have been used in other experiments performed at our laboratory (Chang and Norton, 1996).

Suppression induced by contralateral noise without ipsilateral stimulation also demonstrated an effect which was maximal for $f2$ from 1.4 to 4 kHz, rolling off at higher frequencies. While one possible reason for this rolloff was the high-frequency rolloff above 8 kHz in the contralateral stimulus itself, the fact that the high-frequency tone stimuli also showed a diminishing effect suggests that the rolloff is due to the efferent system being less sensitive to high-frequency stimulation. Studies of guinea pig behavioral thresholds have indicated that the guinea pig ear is relatively flat in sensitivity from 2 to 16 kHz (Fay, 1988), so the differences observed are unlikely to be due to differences in stimulus audibility. A broad effect was observed with 1- and 2-kHz tones, while a narrower effect was observed for the 4-, 8-, and 16-kHz tones. This may be due to differences in efferent innervation patterns between the apical and basal cochlea.

Kirk and Johnstone observed a tonotopic correspondence between the frequency of ipsilateral or contralateral suppressor tones and the primary frequencies, but characterized this specificity to the extent that the effect would be present when the suppressor frequency was within half an octave to the primary frequencies. In our experiments, the maximum suppression occurred when the contralateral suppressor tone was a half-octave below the $f2$ probe tone frequency. While both of our results confirmed previous physiological findings demonstrating the tuning specificity of single olivocochlear fibers and anatomical findings demonstrating the limited spiral extent of arborization (Cody and Johnstone, 1982a; Robertson, 1984, 1985; Robertson and Gummer, 1985; Liberman and Brown, 1986; Brown, 1989), the further finding of a half-octave shift is particularly intriguing, since this phenomenon is also observed following noise exposure (Cody and Johnstone, 1981; Lonsbury-Martin and Meikle, 1978; Chang and Norton, 1995). The presence

of a half-octave shift in response to contralateral suppressor tones is consistent with findings by Guinan and associates (1984) that olivocochlear projections to the contralateral cochlea are generally more basal in location than projections to the ipsilateral cochlea. In Cody and Johnstone's (1982a) data, tuning curves from spiral ganglion afferent neurons recorded from the same site as the recorded efferent neurons show a half-octave shift for four of the eight units presented in the paper.

Ipsilateral primary tones also show a frequency-specific effect. This effect is also seen with addition of the contralateral BBN, as demonstrated in Fig. 7. It was not possible to isolate the specific tuning of the ipsilateral stimulus in this study, since stimulus *pairs* were given. Of note, however, the louder of the two stimuli ($f1$) was 10 dB greater in amplitude, and with the 1.26 $f2/f1$ ratio used in the study, was one-third of an octave below the $f2$ region where evidence indicates that a nonlinear interaction between the $f1$ and $f2$ traveling waves generates the DPOAE (Brown and Kemp, 1984; Brown *et al.*, 1992; Allen and Fahey, 1993). Thus while the ipsilateral stimuli used in this study could not confirm a definite half-octave tuning effect, the results suggested that a half-octave shift *could* be present. A future study utilizing a single ipsilateral suppressor tone, with appropriate stimulus pairs monitoring recovery from suppression, is being planned to investigate whether a difference in tuning between ipsilateral and contralateral efferent effects exists, as suggested by general efferent innervation patterns (Guinan *et al.*, 1984).

Kirk and Johnstone were unable to replicate the suppression of $2f1 - f2$ with contralateral stimulation as demonstrated in previous studies in the guinea pig (Puel and Rebillard, 1990; Kujawa *et al.*, 1993). Both the Puel and the Kujawa experiments utilized Urethane for their anesthetic preparation while the currently described experiments and the Kirk and Johnstone experiments used a ketamine/xylazine mixture to anesthetize the subjects. We were able to consistently demonstrate a suppression effect on $2f - f2$ once the anesthesia wore off, and, furthermore, implicated xylazine as the causative factor. Kirk and Johnstone used a dose of xylazine (50 mg/kg) five times higher than the concentration used in this study (10 mg/kg), which may explain why they were unable to find the effect. Several other investigators have reported that efferent effects can be influenced by anesthesia (Robertson and Gummer, 1985; Brown, 1989). In the Kujawa experiments, when Nembutal was used instead of Urethane, the amount of suppression was reduced from 2.14 to 0.57 dB.

B. Immediate versus delayed effects

Comparing the data from ''light'' versus deeply anesthesized animals indicates that there are two separate types of effects caused by contralateral acoustic stimulation. One is an ''immediate'' effect dependent on the type of anesthetic preparation, while the other is a ''delayed'' effect, requiring several minutes of continuous stimulation to accumulate, and is present even during deep anesthesia. It was not possible to study the immediate effect extensively in our experimental preparation, due to the lightness of anesthesia required in

order to elicit the effect. However, the data obtained did indicate that the changes in $f2-f1$ are of much greater magnitude than the changes in $2f1-f2$. $f2-f1$ was not monitored by Puel and Rebillard or by Kujawa in their experiments. It is surmised that they would have found a much larger effect on $f2-f1$. An additional finding that was not explored in the current study is the initial sharp depression in $f2-f1$ observed with the immediate effect. By focusing upon DPOAE measurements taken in the first second of contralateral stimulation, the apparent magnitude of contralateral suppression may be increased even further.

A valid question to raise at this point is whether the observed ''immediate'' and ''delayed'' effects are truly mediated by the efferent system. Alternative explanations which have been proposed to account for these effects include transcranial crossover of the contralateral sound stimulus and activation of the middle ear muscle reflex. In the experiments described here, careful measurements performed at the ipsilateral ear failed to detect either the contralateral noise or tone stimuli. While it is conceivable that some extremely low unmeasurable level of stimulus crossing over to the ipsilateral ear could affect the cochlear mechanics to some degree, one must note that the ''delayed'' suppression effects elicited by the contralateral stimuli were as large or larger than the effects elicited by continuous ipsilateral stimuli. One would be hard put to explain how such a low crossover signal would provide as large a magnitude of effect as a much higher ipsilateral signal. The cumulative effect over time of ''delayed'' suppression observed in these experiments are not consistent with the middle ear muscle reflex. Since the subjects of these experiments did not have their middle ear muscles transected, one cannot rule out the middle ear muscle reflex as a cause of the ''immediate'' suppression observed in guinea pigs under light anesthesia. Experiments performed by two other groups (Puel and Rebillard, 1990; Kujawa *et al.*, 1993) in guinea pigs which did have their middle ear muscles transected, however, show essentially the same magnitude of contralateral suppression observed here.

Why does anesthesia with xylazine eliminate the ''immediate'' suppression effect? Inquiry into the neurochemistry of the ketamine/xylazine anesthetic agents used for these experiments may provide some answers. Ketamine is an arylcycloalkylamine which induces a dissociative anesthesia by selectively interrupting associative pathways of the brain before producing somesthetic sensory blockade. Poor muscle relaxation is achieved by ketamine alone. Consequently, xylazine is added as an adjuvant. Work on the pharmacological effects of xylazine on guinea pig ileum has indicated that xylazine acts as an $\alpha2$-adrenoceptor agonist (operating specifically at the $\alpha2_a$ subtype) which inhibits neurotensin-induced release of acetylcholine (Blandizzi *et al.*, 1991; Rakovska, 1993). Several researchers have suggested that the contralateral acoustic suppression effect produced by the efferents is mediated by cholinergic receptors of the MOC system (Puel and Rebillard, 1990; Kujawa *et al.*, 1993). The elimination of this effect by xylazine anesthesia is consistent with an acetylcholine-mediated efferent system.

C. Relating the QDT and CDT to OHC receptor potentials, and implications towards regulation of the cochlear amplifier by the efferent system

In the discussion of their results, Kirk and Johnstone attempt to correlate their findings with the anatomy and the neurochemistry of a GABA-ergic mediated LOC efferent system. The actual physiological, or functional significance of the observed modulation of $f2-f1$ remains speculative. They hypothesize that the observed changes in $f2-f1$ reflect a change in the operating bias of the OHC transfer function, possibly mediated by shifts in the resting orientation of OHC stereocilia. The relative stability of $2f1-f2$ suggests that the QDT and the CDT may be generated by two different nonlinearities along the cochlear partition, or at least that a shift in operating point does not change $2f1-f2$ generation.

Further insight into the source of these different nonlinearities can be gained by examining the properties of OHC receptor potentials. Measurements of these responses have indicated that they consist of both cycle-by-cycle (ac) and tonic (dc) components. The ac component results from a symmetrical alternating depolarization and hyperpolarization which may feed back an ac mechanical force upon the basilar membrane. A dc shift is produced as well, due to rectification of the transduction process, resulting in asymmetrical depolarization and hyperpolarization. Dallos indicates that the ac force may be counteracting viscous damping along the cochlear partition while a dc length change in the OHC may be altering the stiffness of the basilar membrane and thus, set the operating point of the micromechanical system (Dallos, 1988, 1992). Analysis of nonlinearities in each of these components reveals that a different type predominates for each component. Hall's (1974) original analysis of an electrical transmission-line model of BM motion showed that even-order distortion products (such as $f2-f1$) are created by introducing an asymmetric relationship between current and resistance,

$$R_i = R_{i0}(1 + \alpha I_i).$$

Odd-order distortion products ($2f1-f2$, $3f1-2f2$, etc.) are created by introducing a symmetric relationship between current and resistance,

$$R_i = R_{i0}(1 + \beta I_i^2).$$

In a model with both symmetric and asymmetric nonlinearities,

$$R_i = R_{i0}(1 + \alpha I_i + \beta I_i^2),$$

odd-order distortion products reflect the symmetric (i.e., ac) nonlinearity while even-order distortion products reflect the asymmetric (i.e., dc) nonlinearity. Dallos and associates (1993) have likewise indicated that mathematical analysis of OHC electromotility shows that the dc distortion component tends to relate to the quadratic component of a polynomial approximation of the nonlinearity.

Several researchers have suggested a feedback mechanism operating within the OHC actively minimizes the dc potential in order to achieve maximum efficiency of electromechanical transduction, optimizing the resulting ac gain of

the system and thus the sensitivity of IHCs (Russell and Kössl, 1991, 1992; Evans *et al.*, 1991; Dallos *et al.*, 1993). The voltage responses of IHCs to tones are increased by negative current (i.e., minimization of the dc potential) and decreased by positive current (Russell and Kössl, 1991). Measures of voltage to OHC length changes have indicated that negative current results in contraction while positive current results in elongation (Evans *et al.*, 1991). A feedback loop between dc receptor potential and OHC length in which the dc is maintained at a low level maximizes the sensitivity of the system. In his studies of basilar membrane (BM) motion in guinea pigs, LePage (1989, 1990) proposed a ''push-pull'' hypothesis in which two separate force-generating mechanisms, having different time courses and working in opposition, control the operating point of the inner hair cells and damping of the partition. LePage (1987) showed that for fixed frequencies the polarity of the mean position of the BM varies systematically with sound levels and place and summates with time since onset, and further suggested that this mean position may be maintained close to some optimal transverse position as regulated by efferent activity.

Nonlinearities within the dc potential manifest themselves in even-order distortion products such as the QDT. Recordings of the QDT at the round window (Nuttall and Dolan, 1993) and in the ear canal using low stimulus levels and different frequency separations (Brown, 1993) have suggested that the QDT originates from a mechanical response of OHCs. Brown (1994) has demonstrated how modulating the even-order distortion with the sum of the stimulus waveforms can generate odd-order distortion. If odd-order distortions are assumed to originate in the ac component of hair cell responses, mechanical even-order distortion produced by OHCs can conceivably result in changes in the ac gain.

Thus, converging lines of inquiry suggest that the quadratic nonlinearity is associated with dc shifts in the operating bias of OHCs. If changes in the acoustic QDT are indeed representative of changes in OHC operating bias, we believe that this study provides strong evidence for efferent regulation of this system. This is best demonstrated in the experiments utilizing extended stimulation. With ipsilateral stimulation only, the QDT was observed to eventually stabilize to a relatively constant level. This stabilization, or ''asymptote,'' has been referred to in previous studies (Whitehead *et al.*, 1991a) and can also be seen in the original data presented by Brown (1988). A possible hypothesis for its significance is that over time, a feedback mechanism within the OHC between dc potentials and cell length changes has optimized the system to the presented continuous primaries. Extended contralateral stimulation seems to decouple the feedback mechanism, resulting in stabilization of the QDT. Electrical biasing of the OHC results in a response in OHC length that is qualitatively similar to the distortion product data presented in this paper. Evans *et al.* (1991) have made measurements of electromechanical length changes induced in solitary OHCs at low stimulus levels, in which they have observed a nonlinearity, consisting of a dc component superimposed on the ac response. In this manuscript they present a figure (Figure 10 on page 297) describing the ac and dc responses to repetitive presentations of a 62.5-mV peak sinu-soidal command, in which the ac response stays relatively stable while the dc response changes dramatically over time. The predicted change in distortion products resulting from nonlinearities in each of these responses will show a change in the QDT in the face of a stable CDT.

The quadratic nonlinearity may thus be an ideal method to specifically study efferently induced changes on cochlear function. We have demonstrated in this study two separate efferently mediated effects—an immediate as well as a delayed effect. Of the distortion products studied, the QDT was the most sensitive indicator for both of these effects. Of particular note is the delayed effect observed with extended contralateral stimulation, which can be made as large as desired by merely extending the stimulation period.

A considerable amount of interest has focused upon the effect of the efferents in noise damage (Cody and Johnstone, 1982b; Rajan, 1983, 1988a, b, 1990; Rajan and Johnstone, 1988a–c; Hildesheimer *et al.*, 1990; Liberman, 1991; Patuzzi and Thompson, 1991; Takeyama *et al.*, 1993; Robertson and Anderson, 1994). One hypothesis is that efferently mediated changes in cochlear bias can reduce the amount of trauma to the hair cells (LePage, 1992; Rajan, 1992). That we found a half-octave shift in an efferently mediated suppression of $f2-f1$ corresponding to the half-octave shift observed in the noise literature further suggests that the efferent system is somehow implicated in outer hair cell responses to acoustic overstimulation. Patuzzi and Thompson (1991) have demonstrated in guinea pigs the protective effect of a contralateral nontraumatic tone (10 kHz at 80 dB SPL) to an ipsilateral traumatic tone (10 kHz at 110 dB SPL) which was not present after cutting the MOC via a midline sagittal section at the floor of the fourth ventricle. The TTS induced by the traumatic tone was reflected in diminished amplitude of the 200-Hz cochlear microphonic (CM), with the CM decrease for animals protected by the contralateral tone smaller than for control animals that were unprotected or deefferented. A finding of particular interest in light of the current study is that while the deefferented animals displayed the same average drop in CM as the unprotected group, they also had a significantly decreased variability in their responses, a finding in much contrast to the striking variations in response to noise exposure typically observed (Cody and Robertson, 1983; Mills, 1976). They use this finding to suggest that interanimal variations in susceptibility to noise may be due to either differing tonic activity or variation in sensitivity of the efferents between individual animals. In the current study, contralateral stimulation effectively ''decoupled'' whatever mechanism was responsible for stabilizing the QDT. We hypothesize that the protective effect of contralateral stimulation observed in these previous studies reflects a ''decoupling'' of the normal regulatory mechanisms, which would have otherwise shifted the operating point of the cochlea to maximize the sensitivity of the system, thus protecting the hair cells. Deefferented animals lacking this protective mechanism end up all receiving the same stabilized maximized amount of mechanical overstimulation, resulting in the observed decreased variability.

One of the problems with both clinical and basic science studies of contralateral suppression is that an extremely

small effect is typically observed, only on the order of 1 or 2 dB. Berlin and associates (1993a, 1994) have recently suggested that Type I afferent neuron disorders in humans manifest themselves via lack of efferent suppression of the TEOAE. To enhance the sensitivity of their measures, they utilize the method of recording the TEOAE concentrated in the time zone from 8–18 ms. An alternative method for increasing sensitivities to efferently mediated effects may be to study the delayed effect instead of the immediate effect. We have shown that the delayed effect on the QDT can basically be made as large as desired by extending the stimulation time. Further experiments are necessary to determine whether these findings also occur in humans.

Allen, J. B., and Fahey, P. F. (**1993**). ''A second cochlear-frequency map that correlates distortion product and neural tuning measurements,'' J. Acoust. Soc. Am. **94**, 809–816.

Altschuler, R. A., and Fex, J. (**1986**). ''Efferent neurotransmitters,'' in *Neurobiology of Hearing: The Cochlea*, edited by R. A. Altschuler, D. W. Hoffman, and R. P. Bobbin (Raven, New York), pp. 383–396.

Berlin, C. I., Hood, L. J., Cecola, R. P., Jackson, D., and Szabo, P. (**1993a**). ''Does Type I afferent neuron dysfunction reveal itself through lack of efferent suppression?,'' Hearing Res. **65**, 40–50.

Berlin, C. I., Hood, L. J., Hurley, A., and Wen, H. (**1994**). ''Contralateral suppression of otoacoustic emissions: An index of the function of the medial olivocochlear system,'' Otolaryngol. Head Neck Surg. **110**, 3–21.

Berlin, C. I., Hood, L. J., Wen, H., Szabo, P., Cecola, R. P., Rigby, P., and Jackson, D. F. (**1993b**). ''Contralateral suppression of non-linear click-evoked otoacoustic emissions,'' Hearing Res. **71**, 1–11.

Blandizzi, C., Doda, M., Tarkovács, G., Del Tacca, M., and Vizi, E. S. (**1991**). ''Functional evidence that acetylcholine release from Auerbach's plexus of guinea-pig ileum is modulated by alpha 2A-adrenoceptor subtype,'' Eur. J. Pharm. **205**, 311–313.

Bledsoe, S. C., Bobbin, R. P., and Puel, J. L. (**1988**). ''Neurotransmission in the inner ear,'' in *Physiology of the Ear*, edited by A. F. Jahn and J. Santos-Sacchi (Raven, New York), pp. 385–406.

Bonfils, P., and Puel, J.-L. (**1987**). ''Functional properties of the crossed part of the medial olivo-cochlear bundle,'' Hearing Res. **28**, 125–130.

Bonfils, P., Remond, M.-C., and Pujol, R. (**1986**). ''Efferent tracts and cochlear frequency selectivity,'' Hearing Res. **24**, 277–283.

Brown, A. M. (**1987**). ''Acoustic distortion from rodent ears: A combination of responses from rats, guinea pigs and gerbils,'' Hearing Res. **31**, 25–37.

Brown, A. M. (**1988**). ''Continuous low level sound alters cochlear mechanics: An efferent effect?,'' Hearing Res. **34**, 27–38.

Brown, A. M. (**1993**). ''Distortion in the cochlea: Acoustic $f2-f1$ at low stimulus levels,'' Hearing Res. **70**, 160–166.

Brown, A. M. (**1994**). ''Modulation of the hair cell motor: A possible source of odd-order distortion,'' J. Acoust. Soc. Am. **96**, 2210–2215.

Brown, A. M., and Gaskill, S. A. (**1990**). ''Measurement of acoustic distortion reveals underlying similarities between human and rodent mechanical responses,'' J. Acoust. Soc. Am. **88**, 840–849.

Brown, A. M., and Kemp, D. T. (**1984**). ''Suppressibility of the $2f_1-f_2$ stimulated acoustic emissions in gerbil and man,'' Hearing Res. **13**, 29–37.

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, M. C., and Nuttall, A. L. (**1984**). ''Efferent control of cochlear inner hair cell responses in the guinea-pig,'' J. Physiol. (London) **354**, 625–646.

Brown, M. C., Nuttall, A. L., and Masta, R. I. (**1983**). ''Intracellular recordings from cochlear inner hair cells: effects of stimulation of the crossed olivocochlear efferents,'' Science **222**, 69–72.

Brown, M. C. (**1989**). ''Morphology and response properties of single olivocochlear fibers in the guinea pig,'' Hearing Res. **40**, 93–110.

Brown, S. E., and Norton, S. J. (**1990**). ''The effects of contralateral acoustic stimulation on the acoustic distortion product, $2f1-f2$,'' Thirteenth midwinter research meeting Association for Research in Otolaryngology, St. Petersburg Beach, Florida, p. 230.

Buño, W. (**1978**). ''Auditory nerve fiber activity influenced by contralateral ear sound stimulation,'' Exp. Neurol. **59**, 62–74.

Chang, K. W., and Norton, S. J. (**1996**). ''The effects of continuous versus interrupted noise exposures on distortion product otoacoustic emissions in guinea pigs,'' Hearing Res. **96**, 1–12.

Cody, A. R., and Johnstone, B. M. (**1981**). ''Acoustic trauma: Single neuron basis for the 'half-octave shift','' J. Acoust. Soc. Am. **70**, 707–711.

Cody, A. R., and Johnstone, B. M. (**1982a**). ''Acoustically evoked activity of single efferent neurons in the guinea pig cochlea,'' J. Acoust. Soc. Am. **72**, 280–282.

Cody, A. R., and Johnstone, B. M. (**1982b**). ''Temporary threshold shift modified by binaural acoustic stimulation,'' Hearing Res. **6**, 199–205.

Cody, A. R., and Robertson, D. (**1983**). ''Variability of noise-induced damage in the guinea pig cochlea: Electrophysiological and morphological correlates after strictly controlled exposures,'' Hearing Res. **9**, 55–70.

Collet, L., Kemp, D. T., Veuillet, E., Duclaux, R., and Moulin, A. (**1990**). ''Effect of contralateral auditory stimuli on active cochlear micromechanical properties in human subjects,'' Hearing Res. **43**, 251–262.

Collet, L., Veuillet, E., Bene, J., and Morgon, A. (**1992**). ''The effects of contralateral white noise on click-evoked emissions in normal and sensorineural ears,'' Audiology **31**, 1–7.

Dallos, P. (**1988**). ''Cochlear neurobiology: Some key experiments and concepts of the past two decades,'' in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 153–188.

Dallos, P. (**1992**). ''The active cochlea,'' J. Neurosci. **12**, 4575–4585.

Dallos, P., Hallworth, R., and Evans, B. N. (**1993**). ''Theory of electrically driven shape changes of cochlear outer hair cells,'' J. Neurophysiol. **70**, 299–323.

Davis, H. (**1983**). ''An active process in cochlear mechanics,'' Hearing Res. **9**, 79–90.

Desmedt, J. E. (**1962**). ''Auditory-evoked potentials from cochlea to cortex as influenced by activation of the efferent olivo-cochlear bundle,'' J. Acoust. Soc. Am. **34**, 1478–1496.

Dewson, J. H. (**1967**). ''Efferent olivocochlear bundle: Some relationships to noise masking and to stimulus attenuation,'' J. Neurophysiol. **30**, 817–832.

Evans, B. N., Hallworth, R., and Dallos, P. (**1991**). ''Outer hair cell electromotility: The sensitivity and vulnerability of the DC component,'' Hearing Res. **52**, 288–304.

Fay, R. R. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay Associates, Winnetka), pp. 375–378.

Fex, J. (**1962**). ''Auditory activity in centrifugal and centripetal cochlear fibres in cat. A study of a feedback system,'' Acta Physiol. Scand. **55**, 1–68.

Folsom, R. C., and Owsley, R. M. (**1987**). ''N1 action potentials in humans. Influence of simultaneous contralateral stimulation,'' Acta Oto-Laryngol. (Stockh.) **103**, 262–265.

Galambos, R. (**1956**). ''Suppression of auditory nerve activity by stimulation of efferent fibres to the cochlea,'' J. Neurophysiol. **19**, 424–437.

Grose, J. H. (**1983**). ''The effect of contralateral stimulation on spontaneous acoustic emissions,'' J. Acoust. Soc. Am. Suppl. 1 **74**, S38.

Guinan, J. J., and Gifford, M. L. (**1988a**). ''Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory nerve fibers. I. Rate-level functions,'' Hearing Res. **33**, 97–114.

Guinan, J. J., and Gifford, M. L. (**1988b**). ''Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory nerve fibers. III. Tuning curves and thresholds at CF,'' Hearing Res. **37**, 29–46.

Guinan, J. J., Warr, W. B., and Norris, B. E. (**1984**). ''Topographic organization of the olivocochlear projections from the lateral and medial zones of the superior olivary complex,'' J. Comp. Neurol. **226**, 21–27.

Hall, J. L. (**1974**). ''Two-tone distortion products in a nonlinear model of the basilar membrane,'' J. Acoust. Soc. Am. **56**, 1818–1828.

Harrison, W. A., and Burns, E. M. (**1993**). ''Effects of contralateral acoustic stimulation on spontaneous otoacoustic emissions,'' J. Acoust. Soc. Am. **94**, 2649–2658.

Hildesheimer, M., Makai, E., Muchnik, C., and Rubinstein, M. (**1990**). ''The influence of the efferent system on acoustic overstimulation,'' Hearing Res. **43**, 263–268.

Kemp, D. T. (**1978**). ''Stimulated acoustic emissions from within the human auditory system,'' J. Acoust. Soc. Am. **64**, 1386–1391.

Kim, D. O. (**1986**). ''Active and nonlinear cochlear biomechanics and the role of outer-hair-cell subsystem in the mammalian auditory system,'' Hearing Res. **22**, 105–114.

Kirk, D. L., and Johnstone, B. M. (**1993**). ''Modulation of f_2-f_1: Evidence

for a GABA-ergic efferent system in apical cochlea of the guinea pig,'' Hearing Res. **67,** 20–34.

Konishi, T., and Slepian, J. Z. (**1971**). ''Summating potential with electrical stimulation of crossed olivocochlear bundles,'' Science **1971,** 483–484.

Kujawa, S. G., and Glattke, T. J. (**1989**). ''Influence of contralateral acoustic stimulation on spontaneous otoacoustic emissions,'' ASHA October, p. 123.

Kujawa, S. G., and Glattke, T. J. (**1990**). ''Influence of contralateral stimulus bandwidth on spontaneous otoacoustic emissions,'' ASHA October p. 156.

Kujawa, S. G., Glattke, T. J., Fallon, M., and Bobbin, R. P. (**1993**). ''Contralateral sound suppresses distortion product otoacoustic emissions through cholinergic mechanisms,'' Hearing Res. **68,** 97–106.

LePage, E. L. (**1987**). ''Frequency-dependent self-induced bias of the basilar membrane and its potential for controlling sensitivity and tuning in the mammalian cochlea,'' J. Acoust. Soc. Am. **82,** 139–154.

LePage, E. L. (**1989**). ''Functional role of the olivo-cochlear bundle: A motor unit control system in the mammalian cochlea,'' Hearing Res. **38,** 177–198.

LePage, E. L. (**1990**). ''Helmholtz revisited: direct mechanical data suggest a physical model for dynamic control of mapping frequency to place along the cochlear partition,'' in *The Mechanics and Biophysics of Hearing,* edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, New York), pp. 278–287.

LePage, E. L. (**1992**). ''Hysteresis in cochlear mechanics and a model for variability in noise-induced hearing loss,'' in *Noise-Induced Hearing Loss,* edited by A. L. Dancer, D. Henderson, R. J. Salvi, and R. P. Hamernik (Mosby-Year Book, St. Louis), pp. 106–115.

Liberman, M. C. (**1989**). ''Rapid assessment of sound-evoked olivocochlear feedback: Suppression of compound action potentials by contralateral sound,'' Hearing Res. **38,** 47–56.

Liberman, M. C. (**1990**). ''Effects of chronic cochlear de-efferentation on auditory-nerve response,'' Hearing Res. **49,** 209–224.

Liberman, M. C. (**1991**). ''The olivocochlear efferent bundle and susceptibility of the inner ear to acoustic injury,'' J. Neurophysiol. **65,** 123–132.

Liberman, M. C., and Brown, M. C. (**1986**). ''Physiology and anatomy of single olivocochlear neurons in the cat,'' Hearing Res. **24,** 17–36.

Littman, T. A., Cullen, Jr., J. K., and Bobbin, R. P. (**1992**). ''The effect of olivocochlear bundle transection on tuning curves and acoustic distortion products,'' J. Acoust. Soc. Am. **92,** 1945–1962.

Lonsbury-Martin, B. L., and Meikle, M. B. (**1978**). ''Neural correlates of auditory fatigue. Frequency dependent changes in activity of single cochlear nerve fibers,'' J. Neurophysiol. **41,** 987–1006.

Mills, D. M., and Rubel, E. W. (**1994**). ''Variation of distortion product otoacoustic emissions with furosemide injection,'' Hearing Res. **77,** 183–199.

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, J. H. (**1976**). ''Individual differences in noise-induced hearing loss,'' in *Hearing and Davis: Essays Honoring Halowell Davis,* edited by S. K. Hirsch, D. H. Eldredge, J. J. Hirsch, and S. R. Silverman (Washington U.P., St. Louis), pp. 73–84.

Mott, J. B., Norton, S. J., Neely, S. T., and Warr, W. B. (**1989**). ''Changes in spontaneous otoacoustic emissions produced by acoustic stimulation of the contralateral ear,'' Hearing Res. **38,** 229–242.

Moulin, A., Collet, L., and Duclaux, R. (**1993**). ''Contralateral auditory stimulation alters acoustic distortion products in humans,'' Hearing Res. **65,** 193–210.

Moulin, A., Collet, L., and Morgon, A. (**1992**). ''Influence of spontaneous otoacoustic emissions (SOAE) on acoustic distortion product input/output functions: Does the medial efferent system act differently in the vicinity of an SOAE?,'' Acta Otolaryngol. (Stockh.) **112,** 210–214.

Mountain, D. C. (**1980**). ''Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alter cochlear mechanics,'' Science **210,** 71–72.

Murata, K., Tanahashi, T., Horikawa, J., and Funai, H. (**1980**). ''Mechanical and neural interactions between binaurally applied sounds in cat cochlear nerve fibers,'' Neurosci. Lett. **18,** 289–294.

Nieder, P., and Nieder, I. (**1970a**). ''Stimulation of efferent olivocochlear bundle causes release from low level masking,'' Nature (London) **227,** 184–185.

Nieder, P. C., and Nieder, I. (**1970b**). ''Crossed olivocochlear bundle: Electrical stimulation enhances masked neural responses to loud clicks,'' Brain Res. **21,** 135–137.

Nuttall, A. L., and Dolan, D. F. (**1993**). ''Intermodulation distortion ($F_2 - F_1$) in inner hair cell and basilar membrane responses,'' J. Acoust. Soc. Am. **93,** 2061–2068.

Patuzzi, R. B., and Thompson, M. L. (**1991**). ''Cochlear efferent neurones and protection against acoustic trauma: Protection of outer hair cell receptor current and interanimal variability,'' Hearing Res. **54,** 45–58.

Puel, J.-L., and Rebillard, G. (**1990**). ''Effect of contralateral sound stimulation on the distortion product $2f_1 - f_2$: Evidence that the medial efferent system is involved,'' J. Acoust. Soc. Am. **87,** 1630–1635.

Pujol, R., and Lenoir, M. (**1986**). ''The four types of synapses in the organ of Corti,'' in *Neurobiology of Hearing: The Cochlea,* edited by R. A. Altschuler, R. P. Bobbin, and D. W. Hoffman (Raven, New York), pp. 161–172.

Rabinowitz, W. M., and Widin, G. P. (**1984**). ''Interaction of spontaneous otoacoustic emissions and external sounds,'' J. Acoust. Soc. Am. **76,** 1713–1720.

Rajan, R. (**1988a**). ''Effect of electrical stimulation of the crossed olivocochlear bundle on temporary threshold shifts in auditory sensitivity. I. Dependence on electrical stimulation parameters,'' J. Neurophysiol. **60,** 549–568.

Rajan, R. (**1988b**). ''Effect of electrical stimulation of the crossed olivocochlear bundle on temporary threshold shifts in auditory sensitivity. II. Dependence on the level of temporary threshold shifts,'' J. Neurophysiol. **60,** 569–579.

Rajan, R. (**1990**). ''Electrical stimulation of the inferior colliculus at low rates protects the cochlea from auditory desensitization,'' Brain Res. **506,** 192–204.

Rajan, R. (**1992**). ''Protective functions of the efferent pathways to the mammalian cochlea: A review,'' in *Noise-Induced Hearing Loss,* edited by A. L. Dancer, D. Henderson, R. J. Salvi, and R. P. Hamernik (Mosby–Year Book, St. Louis), pp. 429–444.

Rajan, R., and Johnstone, B. M. (**1983**). ''Crossed cochlear influences on monaural temporary shifts,'' Hearing Res. **9,** 279–274.

Rajan, R., and Johnstone, B. M. (**1988a**). ''Electrical stimulation of cochlear efferents at the round window reduces auditory desensitization in guinea pigs. I. Dependence on electrical stimulation parameters,'' Hearing Res. **36,** 53–74.

Rajan, R., and Johnstone, B. M. (**1988b**). ''Electrical stimulation of cochlear efferent at the round window reduces auditory desensitization in guinea pigs. II. Dependence on level of temporary threshold shifts,'' Hearing Res. **36,** 75–88.

Rajan, R., and Johnstone, B. M. (**1988c**). ''Binaural acoustic stimulation exercises protective effects at the cochlea that mimic the effects of electrical stimulation of an auditory efferent pathway,'' Brain Res. **459,** 241–255.

Rajan, R., Robertson, D., and Johnstone, B. M. (**1990**). ''Absence of tonic activity of the crossed olivocochlear bundle in determining compound action potential thresholds, amplitudes and masking phenomena in anaesthetised guinea pigs with normal hearing sensitivities,'' Hearing Res. **44,** 195–208.

Rakovska, A. D. (**1993**). ''Functional and neurochemical evidence that neurotensin-induced release of acetylcholine from Auerbach's plexus of guinea-pig ileum is presynaptically controlled via alpha 2-adrenoceptors,'' Neurochem. Res. **18,** 737–741.

Rasmussen, G. L. (**1946**). ''The olivary peduncle and other fiber projections of the superior olivary complex,'' J. Comp. Neurol. **84,** 141–220.

Robertson, D. (**1984**). ''Horseradish peroxidase injection of physiologically characterised afferent and efferent neurones in the guinea pig spiral ganglion,'' Hearing Res. **15,** 113–121.

Robertson, D. (**1985**). ''Brainstem location of efferent neurones projecting to the guinea pig cochlea,'' Hearing Res. **20,** 79–84.

Robertson, D., and Anderson, C. J. (**1994**). ''Acute and chronic effects of unilateral elimination of auditory nerve activity on susceptibility to temporary deafness induced by loud sound in the guinea pig,'' Brain Res. **646,** 37–43.

Robertson, D., and Gummer, M. (**1985**). ''Physiological and morphological characterization of efferent neurones in the guinea pig cochlea,'' Hearing Res. **20,** 63–77.

Rossi, G., Actis, R., Solero, P., Rolando, M., and Pejrone, M. D. (**1993**). ''Cochlear interdependence and micromechanics in Man and their relations with the activity of the medial olivocochlear efferent system (MOES),'' J. Laryngol. Otol. **107,** 883–891.

Russell, I. J., and Kössl, M. (**1991**). ''The voltage responses of hair cells in the basal turn of the guinea-pig cochlea,'' J. Physiol. (London) **435,** 493–511.

Russell, I. J., and Kössl, M. (**1992**). ''Voltage responses of outer hair cells in the basal turn of the guinea-pig cochlea: significance for electromotility and desensitization,'' Proc. R. Soc. Lond. B. Biol. Sci. **247**, 97–105.

Ryan, S., Kemp, D. T., and Hinchliffe, R. (**1991**). ''The influence of contralateral acoustic stimulation on click-evoked otoacoustic emissions in humans,'' Br. J. Audiol. **25**, 391–397.

Schloth, E., and Zwicker, E. (**1983**). ''Mechanical and acoustical influences on spontaneous oto-acoustic emission,'' Hearing Res. **11**, 285–293.

Siegel, J. H., and Kim, D. O. (**1982**). ''Efferent neural control of cochlear mechanics? Olivocochlear bundle stimulation affects cochlear biomechanical nonlinearity,'' Hearing Res. **6**, 171–182.

Takeyama, M., Kusakari, J., Nishikawa, N., and Wada T. (**1992**). ''The effect of crossed olivo-cochlear bundle stimulation on acoustic trauma,'' Acta Otolaryngol. (Stockh.) **112**, 205–209.

Teas, D. C., Konishi, T., and Nielson, D. W. (**1972**). ''Electrophysiological studies on the spatial distribution of the crossed olivocochlear bundle along the guinea pig cochlea,'' J. Acoust. Soc. Am. **51**, 1256–1264.

Terayama, Y., and Yamamoto, K. (**1971**). ''Olivo-choclear bundle in the guinea pig cochlea after central transection of the crossed bundle,'' Acta Otolaryngol. (Stock.) **72**, 385–396.

Terayama, Y., Yamamoto, K., and Sakamoto, T. (**1969**). ''The efferent olivo-chochlear bundle in the guinea pig cochlea,'' Ann. Otol. Rhinol. Laryngol. **78**, 1254–1268.

Veuillet, E., Collet, L., and Duclaux, R. (**1991**). ''Effect of contralateral acoustic stimulation on active cochlear micromechanical properties in human subjects: Dependence on stimulus variables,'' J. Neurophysiol. **65**, 724–735.

Veuillet, E., Collet, L., and Morgon, A. (**1992**). ''Differential effects of ear-canal pressure and contralateral acoustic stimulation on evoked otoacoustic emissions in humans,'' Hearing Res. **61**, 47–55.

Warr, W. B., and Guinan, J. J. (**1978**). ''Efferent innervation of the organ of Corti: two separate systems,'' Brain Res. **173,** 152–155.

Warr, W. B., Guinan, J. J., and White, J. S. (**1986**). ''Organization of the efferent fibers: The lateral and medial olivocochlear systems,'' in *Neurobiology of Hearing: The Cochlea*, edited by R. A. Altschuler, D. W. Hoffman, and R. P. Bobbin (Raven, New York), pp. 333–348.

Warren, E. H., and Liberman, M. C. (**1989a**). ''Effects of contralateral sound on auditory-nerve responses. I. Contributions of cochlear efferents,'' Hearing Res. **37**, 89–104.

Warren, E. H., and Liberman, M. C. (**1989b**). ''Effects of contralateral sound on auditory-nerve responses. II. Dependence on stimulus variables,'' Hearing Res. **37**, 105–122.

Whitehead, M. L., Lonsbury-Martin, B. L., and Martin, G. K. (**1991a**). ''Slow variation of the amplitude of acoustic distortion at $f_2 - f_1$ in awake rabbits,'' Hearing Res. **51**, 293–300.

Whitehead, M. L., Martin, G. K., and Lonsbury-Martin, B. L. (**1991b**). ''Effects of the crossed acoustic reflex on distortion-product otoacoustic emission in awake rabbits,'' Hearing Res. **51**, 55–72.

Wiederhold, M. L. (**1986**). ''Physiology of the olivocochlear system,'' in *Neurobiology of Hearing: The Cochlea,* edited by R. A. Altschuler, D. W. Hoffman, and R. P. Bobbin (Raven, New York), pp. 349–370.

Wiederhold, M. L., and Kiang, N. Y. S. (**1970**). ''Effects of electrical stimulation of the crossed olivocochlear bundle on single auditory-nerve fibers in the cat,'' J. Acoust. Soc. Am. **48**, 950–965.

Williams, E. A., Brookes, G. B., and Prasher, D. K. (**1993**). ''Effects of contralateral acoustic stimulation on otoacoustic emissions following vestibular neurectomy,'' Scand. Audiol. **22,** 197–203.

Williams, E. A., Brookes, G. B., and Prasher, D. K. (**1994**). ''Effects of olivocochlear bundle section on otoacoustic emissions in humans: efferent effects in comparison with control subjects,'' Acta Oto-Laryngol. **114,** 121–129.

Medial efferent inhibition suppresses basilar membrane responses to near characteristic frequency tones of moderate to high intensities

Ian J. Russell and Euan Murugasu[a)]

School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, United Kingdom

Efferent inhibition of low and medium speontaneous-rate fiber activity in response to characteristic frequency (CF) tones is greater at levels between 50 and 75 dB SPL than at lower levels, and even occurs at levels of 100 dB SPL [Guinan and Stankovic, J. Acoust. Soc. Am. **100**, 1680–1690 (1996)]. The self-mixing effect of a laser diode was used to measure tone-evoked, basilar membrane (BM) displacements in the basal turn of the guinea pig cochlea during electrical stimulation of the medial efferent system. Efferent stimulation suppresses BM displacement and, for frequencies close to the CF of the measurement site and above, attenuation of tone-evoked displacements appears to be greatest for levels between 50 and 75 dB SPL, and efferent attenuation could exceed 10 dB SPL for tones at 90 dB SPL. By comparison, the attenuation of the BM vibrations caused by tones at frequencies below CF was found to be greatest for low sound pressure levels. The findings of Guinan and Stankovic therefore have their origin in events present in the vibrations of the BM. © *1997 Acoustical Society of America.* [S0001-4966(97)01909-7]

PACS numbers: 43.64.Kc, 43.64.Ld, 43.64.Me [RDF]

INTRODUCTION

The sensitivity of the cochlea is controlled at the level of the outer hair cells (OHCs) by fibers of the medial efferent system. These originate from medial neurones in the olivary complex of the brain stem (see Guinan, 1996 for a recent review). The post-synaptic action of the efferent fibers is frequency selective. Maximum inhibition in response to activation of the efferent system is observed close to the tips of isoresponse tuning curves derived from basilar membrane (BM) vibrations (Dolan and Nuttall, 1994; Murugasu and Russell, 1996a), inner hair cell receptor potentials (Brown and Nuttall, 1984), and afferent fiber activity (Wiederhold and Kiang, 1970). Studies on isolated OHCs (Housley and Ashmore, 1991; Blanchet *et al.*, 1996; Evans, 1996), gross cochlear potentials (Sridhar *et al.*, 1995a), and BM displacements recorded *in situ* (Murugasu and Russell, 1996b) have shown that the postsynaptic action of the putative efferent transmitter, acetylcholine (ACh) (see Eybalin, 1993 for a review), is calcium dependent. It is currently understood that, during efferent stimulation and the release of ACh from the presynaptic terminal, there follows an influx of calcium through the ACh-ligand gated channels of the OHC post-synaptic membrane. The post-synaptic rise in calcium either directly, or indirectly, through a ''slow'' medial efferent effect (Sridhar *et al.*, 1995a,b; Murugasu and Russell, 1996b; Sridhar *et al.*, 1997), leads to hyperpolarization of the OHC membrane potential through the gating of calcium activated potassium channels distributed in their basolateral membranes. However, how the efferent fibers exert their effects on the mechanical properties of the cochlear partition (Murugasu and Russell, 1996a; Dallos *et al.*, 1997), and their role in auditory perception (see Guinan, 1996) remains enigmatic.

Although it was originally thought that efferent inhibition is most effective at low sound levels, Guinan and Stankovic (1996) have recently revealed that this might be a misconception. They recorded from fibers in the cat auditory nerve and found that for levels below 50 dB SPL medial efferent inhibition was similar for all fibers with similar CFs. However, for tone levels between 50 and 75 dB SPL, low spontaneous-rate fibers and medium spontaneous-rate fibers often showed much larger inhibition than at low levels, and some inhibition was even present at 100 dB SPL. We have used the self-mixing effect of a laser diode (Kössl and Russell, 1995) to measure tone-evoked, BM displacements in the basal turn of the guinea pig cochlea to see if this level-dependent inhibition is manifested in the vibrations of the BM.

I. METHODS

A detailed description of the methods of the preparation, experimental techniques, and stimulus presentation has been published (Murugasu and Russell, 1996a). Recordings were made from pigmented guinea pigs (230–320 g) anaesthetized with the neurolept anaesthetic technique (0.06-mg/kg atropine sulphate s.d., 30-mg/kg pentobarbitone i.p., 4-mg/kg droperidol i.m.; 1-mg/kg phenoperidine i.m.) (Evans, 1979). The animals were tracheotomized, artificially respired, and core temperatures were maintained at 37 °C. The compound action potential (CAP) was recorded through a Ag/AgCl electrode placed on the round window. A CAP audiogram was obtained over the range 1–30 kHz before and after each major surgical procedure and after each series of BM displacement measurements made at a single location on the BM.

The medial efferents were stimulated electrically at the point where their fibers cross the floor of the 4th ventricle, close to the genu of the facial nerve. The fibers were shocked through a pair of flexible, teflon coated wire electrodes which were advanced rostrally along the floor of the 4th

ventricle until the tips were close to the point of crossing of the medial efferent fibres. The position of the electrodes was adjusted so that a train of 16, 0.2-ms pulses at 200/s, and with amplitudes of 1–2 V caused only the tips of the pinna to move. The stimulus amplitude required to excite the medial efferents was usually less than that required to cause muscle contraction and usually did not exceed 1.5 V. Since the impedance of the electrodes, measured *in situ*, was about 0.5 MOhms, the current necessary to excite the OCB was only about 3 μA. Through careful placement of the electrodes and limitation of current strength, it was possible to avoid contraction of the middle ear muscles and, hence, the large displacements of the scala tympani perilymph which are associated with these movements. Following the successful placement of the OCB stimulating electrodes, the animals were paralysed with Flaxedil (gallamine triethiodide) (0.4 mg/kg) and artificially respired with carbogen (95% oxygen, 5% carbon dioxide).

BM displacements were measured using the self-mixing effect of a laser diode. This technique is sensitive enough to measure light directly reflected from the BM. Speed of measurement was a crucial consideration in these experiments so most data were obtained without averaging. The interferometer formed a spot on the BM 10 μm in diameter with a <5-μm depth of focus at a distance of 5 cm from the front surface of the focusing lens. Self-mixing in a laser diode involves coupling back a small part of the laser light into the laser cavity. Mixing of this light with the light produced in the cavity is obtained when the target is displaced along the beam axis. The intensity of the laser varies sinusoidally with a period corresponding to a displacement of $\lambda/2$, where λ is the wavelength of the laser light (670 nm). Provided the interferometer is operating about a resting position from the target which is an integer function of $\lambda/2$, the method underestimates the displacement by <3% for changes in displacement of $<\pm45$ nm about the resting position. The displacement-dependent effect was measured in the signal of the photodiode which is located behind the laser crystal in the diode laser housing. The bandwidth of the displacement-dependent signal was 200 kHz and the equivalent noise floor is <10 pm. Calibration of the signal was achieved *in situ*, and with respect to the measurement point on the BM, by displacing the interferometer a known amount by a piezoelectric driver. Voltage responses from the interferometer proportional to BM displacement were analyzed with a two channel lock-in amplifier (EG & G, Princeton, Model 510) set in quadrature. The in-phase and quadrature outputs were digitized and the phase and the magnitude were computed.

Sound was delivered to the tympanic membrane by a closed acoustic system comprising a Bruel & Kjaer 4133 1/2-in. microphone with 1/4-in. coupler for delivering tones and a Bruel and Kjaer 3134 1/2-in. microphone with a 1/4-in. coupler for monitoring the sound pressure at the tympanum. Pure-tone stimuli, 40-ms duration with 2-ms cosine-shaped rise and fall times, delivered every 200 ms, were generated by a Hewlett Packard 325 synthesiser coupled to a computer-controlled gate and an attenuator. The order in which stimulus levels were presented was not randomised because hair cell responses become desensitised only at the highest stimu-
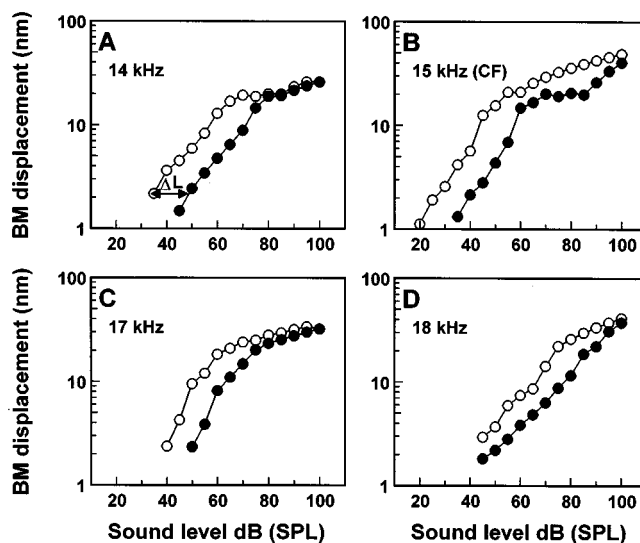
FIG. 1. BM displacement as functions of sound level measured, without (open circles) and with (filled circles) efferent stimulation, from a single position on the BM with a CF of 15 kHz for four different frequencies centered around the CF. The efferent-induced level shift (ΔL) is indicated in the 14-kHz figure and plotted as a function of sound level in Fig. 2.

lus levels (Cody and Russell, 1985, 1995). The sound system was calibrated *in situ* for frequencies between 1 and 30 kHz such that known sound pressure levels expressed in dB SPL (dB *re*: 2×10^{-5} Pa) were used in the experiments.

II. RESULTS

We have used the method developed by Guinan and Stankovic (1996) to obtain a measure of the attenuation, or level shift (ΔL) of tone-elicited BM displacements following electrical stimulation of the medial efferent fibers. They have defined ΔL as the amount by which the sound level must be increased with efferent stimulation to produce the same response as that without efferent stimulation. Stimulation of the medial efferent system caused a lateral shift of the fast rising phase of the level function toward higher sound levels. An example of this shift is shown in Figs. 1A and 2A for level functions derived from measurements of BM displacements in response to a 14-kHz tone which was below the 15-kHz characteristic frequency (CF) of the measurement place (Figs. 1B and 2B). The shift toward higher SPL was largely due to a reduction in the dynamic range of the compressive region of the level function. Thus for the particular conditions of the measurement illustrated in Fig. 1A (14 kHz) (and also in those shown in Figs. 3A and B, and 4A and B for measurements made at frequencies below that of the 16 kHz CF in another preparation) the greatest ΔL occurred at the lowest sound pressure levels while, for tone levels of 80 dB SPL, the level shift was less than 5 dB SPL (Figs. 2A and 4A and B).

ΔL was found to depend not only on tone level but also on the frequency of the tone. For a tone close to and above the CF of the recording site on the BM (Fig. 1B–D and Fig. 3C–F), efferent stimulation resulted in a parallel shift of the steep portion of the curve associated with a strong reduction in the dynamic range of the compression of the level function. This was accompanied with a decrease in the magnitude and an increase in the slope of the compression, which is
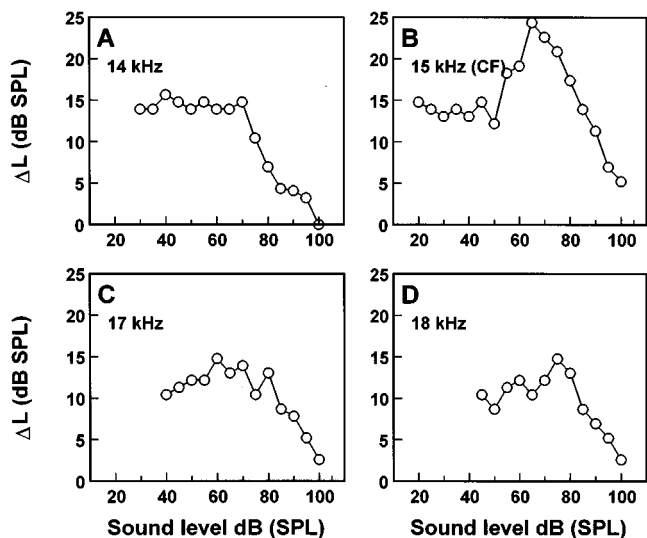
FIG. 2. The level shift (ΔL) as a function of sound level derived from the displacement/level functions shown in Fig. 1.

most noticable for frequencies well above the CF (Fig. 1D, and 3E and F). Furthermore there was a more gradual transition between the steep and compressive regions of the level functions, particularly at the CF (Figs. 1B and 3C). The consequence of these changes in compression on ΔL was to produce maximum attenuation for levels between 50 and 75

FIG. 4. The level shift (ΔL) as a function of sound level derived from the displacement/level functions shown in Fig. 3.

dB SPL and more than 10 dB of attenuation at 85 dB SPL for frequencies close to CF (Figs. 2B and 4C and D). The level shift of BM responses to tones at the greatest intensities employed in these experiments was estimated according to the method of Guinan and Stankovic (1996) by extrapolating the level function measured during efferent stimulation from the slope of the last two points of the level functions. The pattern of the efferent-induced level shift as a function of CF tone level was consistent across preparations. This can be seen in Fig. 5 where the mean ΔL, measured from eight different

FIG. 3. BM displacement as functions of sound level measured, without (open circles) and with (filled circles) efferent stimulation, from a single position on the BM with a CF of 16 kHz for six different frequencies centered around the CF.
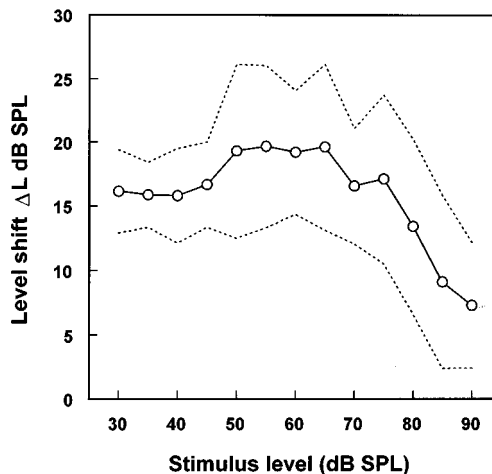
FIG. 5. The mean level shift (ΔL) + standard deviation (dotted lines) as a function of sound level for CF tones derived from eight different preparations.

preparations for tones measured at the CF of the measurement site (14–20 kHz), reached a maximum between 50 and 75 dB SPL and that ΔL approached 10 dB SPL for tone levels of 85 dB SPL.

III. DISCUSSION

We have shown that electrical stimulation of the cochlear efferent system has two predominant effects on the vibrations of the BM to tones at frequencies close to and above the CF. At these frequencies, efferent stimulation causes a lateral shift of the steep portion of the level function to higher intensities and depresses the compressive saturation of BM vibrations. If ΔL is measured for these highly compressive, saturated regions of the level functions, then attenuation of tone-evoked displacements appears to be greatest for levels between 50 and 75 dB SPL and efferent attenuation could even exceed 10 dB SPL for the tones at 90 dB SPL. In contrast, the attenuation of the BM vibrations caused by tones at frequencies below CF was found to be greatest for low sound pressure levels and resulted only in a lateral shift of the level functions.

Our finding for the efferent suppression of BM responses to CF tones is in agreement with those of Guinan and Stankovic (1996), who found that OCB stimulation produced the greatest ΔL, and thus appeared to be most effective in attenuating the responses of low spontaneous and medium spontaneous auditory nerve fibers to moderately loud CF tones. Guinan and Stankovic attributed the effects they observed to events taking place at the IHC afferent synapse. They suggested that the efferent-induced gross extracellular potential (MOC potential), which can be recorded in the cochlea, may reduce the release of afferent transmitter from the presynaptic membrane of the IHC synapse and increase the threshold for neural excitation at the post-synaptic terminal by hyperpolarizing the IHCs and afferent nerve terminals. This speculation remains to be verified. However, it is clear from the results presented here that the attenuation of cochlear responses to moderately loud CF tones originates in events which precede signal processing at the inner hair cell afferent synapse and which are present in the vibrations of the BM.

The similarity in the responses of the BM and those of the low spontaneous and medium spontaneous afferent fibers to efferent stimulation might be because the saturation of firing rate in these two classes of auditory afferent fiber is largely determined by the saturating nonlinear vibration of the BM, while that of the high spontaneous units is probably limited by the saturation of synaptic and spike generating mechanisms at the inner hair cell (IHC) afferent synapse (Yates et al., 1990; Zagaeski et al., 1994). Thus the action of the medial efferents is to cause a decrease in sensitivity across the entire dynamic range of basilar membrane vibration and hence that of the peripheral auditory system, even reducing the level of BM vibrations in the compressive regions of the level function. For low level tones, which are within the response domain of the high spontaneous-rate fibers, the decrease in BM sensitivity might have a role for example, in reducing the masking effects of low-level background noise (Kawase et al., 1993; Guinan, 1996). At the other extent of the dynamic range, within the scope of the low-spontaneous fibers, the effect of medial efferent inhibition would be to extend the dynamic range of auditory sensation above that of moderate to loud background noise. At the highest levels and for sustained activation of the medial efferents for 1–2 min, the efferent "slow" effect (Sridhar et al., 1995a,b) may have a role in reducing the impact of temporary threshold shift (Cody and Johnstone, 1982; Reiter and Liberman, 1995), but does not appear to protect the auditory system against permanent damage (Liberman and Gao, 1995).

The dynamic range of the compressive region of the level function is most extensive for frequencies at and above the CF of the measurement point on the BM. OCB stimulation reduces both the magnitude and dynamic range of this compression (Wiederhold, 1970; Guinan and Gifford, 1988; Brown and Nuttall, 1984; Murugasu and Russell, 1996). However, the increased slope of the compression and the more gradual transition between the steep and compressive regions of the level function largely account for the apparent increase in the OCB-induced level shift of tone-evoked BM displacements for levels between 50 and 75 dB SPL for frequencies at and above the CF of the measurement point on the BM. It was difficult to obtain an accurate measure of ΔL for near CF tone levels for which the BM responses were very compressive (e.g., Fig. 3D). This was because it became necessary to extrapolate between two nearly parallel functions in order to obtain an intersect and hence the value of ΔL (Guinan and Stankovic, 1996). Thus to describe the action of the efferents as causing a lateral shift of the level function at these levels and frequencies is probably misleading. Rather, it may be more accurate to say that efferent activation has an overall suppressive effect on BM mechanics which suppresses the responses of the BM at all levels and frequencies for which efferent stimulation is effective.

It is likely that the efferent stimulation causes a decrease in the forces exerted by the OHCs on the cochlear partition (Murugasu and Russell, 1996a). The compressive, nonlinear vibration of the BM is thought to be determined mainly by the compressive nonlinear properties of the OHC transducer conductance (Patuzzi et al., 1989; Santos-Sacchi, 1993) which, in turn, is believed to control the voltage-dependent OHC motility (Brownell et al., 1985; Zenner, 1986; Ashmore, 1987; Santos-Sacchi and Dilger, 1988; Evans et al., 1989). However, it remains to be seen if changes in the forward or reverse transducer functions or changes in the mechanical properties of the OHCs (Dallos et al., 1997) are the prime cause of the efferent-induced suppression of BM responses to tones.

Ashmore, J. F. (**1987**). ''A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier,'' J. Physiol. (London) **338,** 323–347.

Blanchet, C., Eròstegui, C., Sugasawa, M., and Dulon, D. (**1996**). ''Acetylcholine-induced potassium current of guinea pig outer hair cells: its dependence on a calcium influx through nicotinic-like receptors,'' J. Neurosci. **16,** 2574–2584.

Brown, M. C., and Nuttall, A. L. (**1984**). ''Efferent control of cochlear inner hair cell responses in the guinea pig,'' J. Physiol. (London) **354,** 625–646.

Brownell, W. E., Bader, C. R., Bertrand, D., and Ribaupirre, Y. (**1985**). ''Evoked mechanical responses of isolated cochlear outer hair cells,'' Science **227,** 194–196.

Cody, A. R., and Russell, I. J. (**1985**). ''Outer hair cells in the mammalian cochlea and noise-induced hearing loss,'' Nature (London) **315,** 662–665.

Cody, A. R., and Russell, I. J. (**1995**). ''Time-varying responses of mammalian hair cells to isoamplitude acoustic stimulation,'' Aud. Neurosci. **1,** 351–361.

Dallos, P., He, D. Z. Z., Evans, B. N., and Sziklai, I. (**1997**). ''Efferent control of cochlear mechanics: outer hair cells,'' in *Diversity in Auditory Mechanics*, edited by E. R. Lewis, G. R. Long, P. A. Narins, and C. R. Steel (World Scientific, Singapore) (in press).

Dolan, D. F., and Nuttall, A. L. (**1994**). ''Basilar membrane movement evoked by sound is altered by electrical stimulation of the crossed olivocochlear bundle,'' Assoc. Res. Otolaryngol. Abstr. **17,** 356.

Evans, E. F. (**1979**). ''Neuroleptanaesthesia for the guinea pig, an ideal anaesthetic procedure for long term physiological studies of the cochlea,'' Arch. Otolaryngol. **105,** 185–186.

Evans, B. N., Dallos, P., and Hallworth, R. (**1989**). ''Asymmetries in motile responses of outer hair cells in simulated *in vivo* conditions,'' in *Cochlear Mechanisms: Structure, Function and Models*, edited by D. Kemp and J. P. Wilson (Plenum, New York), pp. 205–206.

Evans, M. G. (**1996**). ''Acetylcholine activates two currents in guinea-pig outer hair cells,'' J. Physiol. (London) **491,** 563–578.

Eybalin M. (**1993**). ''Neurotransmitters and neuromodulators of the mammalian cochlea,'' Physiol. Rev. **73,** 309–373.

Guinan, J. C. (**1996**). ''Physiology of olivocochlear efferents,'' in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer, New York), pp. 435–502.

Guinan J. J., and Gifford, M. L. (**1988**). ''Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory nerve fiber. I. Rate-level functions,'' Hearing Res. **33, 97–114.**

Guinan, J. C., and Stankovic, K. M. (**1996**). ''Medial efferent inhibition produces the largest equivalent attenuation at moderate to high sound levels in cat auditory nerve fibers,'' J. Acoust. Soc. Am. **100,** 1680–1690.

Housley, G. D., and Ashmore, J. F. (**1991**). ''Direct measurement of the action of acetylcholine on isolated outer hair cells of the guinea pig cochlea,'' Proc. R. Soc. London, Ser. B **244,** 161–167.

Kawase, T., Delgutte, B., and Liberman, M. C. (**1993**). ''Anti-masking effects of the olivocochlear reflex, II: Enhancement of the auditory nerve response to masked tones,'' J. Neurophysiol. **70,** 2518–2549.

Kössl, M., and Russell, I. J. (**1995**). ''Basilar membrane resonance in the cochlea of the moustached bat,'' Proc. Natl. Acad. Sci. USA **92,** 276–279.

Liberman, M., and Gao, W.-Y. (**1995**). ''Chronic cochlear de-efferentation and susceptibility to permanent acoustic injury,'' Hearing Res. **90,** 158–168.

Murugasu, E., and Russell, I. J. (**1996a**). ''The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea,'' J. Neurosci. **16,** 325–332.

Murugasu, E., and Russell, I. J. (**1996b**). ''The role of calcium on the effects of intracochlear acetylcholine perfusion on basilar membrane displacement in the basal turn of the guinea pig cochlea,'' Aud. Neurosci. **2,** 363–376.

Patuzzi, R. B., Yates, G. K., and Johnstone, B. M. (**1989**). ''Outer hair cell current and sensorineural hearing loss,'' Hearing Res. **42,** 47–72.

Reiter, E. R., and Liberman, M. C. (**1995**). ''Efferent-mediated protection from acoustic overexposure: Relation to slow effects of olivocochlear stimulation,'' J. Neurophysiol. **73,** 506–514.

Santos-Sacchi, J. (**1993**). ''Harmonics of outer hair cell motility,'' Biophys. J. **65,** 2217–2227.

Santos-Sacchi, J., and Dilger, J. P. (**1988**). ''Whole cell currents and mechanical responses of outer hair cells,'' Hearing Res. **35,** 143–150.

Sridhar, T. S., Liberman, M. C., Brown, M. C., and Sewell, W. F. (**1995a**). ''A novel cholinergic ''slow effect'' of efferent stimulation on cochlear potentials in the guinea pig,'' J. Neurosci. **15,** 3667–3678.

Sridhar, T. S., Brown, M. C., and Sewell, W. F. (**1995b**). ''Molecular mechanisms involved in olivocochlear slow effects,'' ARO Abstracts, XVIIIth Midwinter Meeting, p. 172.

Sridhar, T. S., Brown, M. C., and Sewell, W. F. (**1997**). ''Unique postsynaptic signalling at the hair cell efferent synapse permits calcium to evoke changes on two different time scales,'' J. Neurosci. (in press).

Wiederhold, M. L. (**1970**). ''Variations in the effects of electrical stimulation of the crossed olivocochlear bundle on cat single auditory nerve-fibre responses to tone bursts,'' J. Acoust. Soc. Am. **48,** 966–977.

Wiederhold, M. L., and Kiang, N. Y.-S. (**1970**). ''Effects of electrical stimulation of the crossed olivocochlear bundle on single auditory-nerve fibres in the cat,'' J. Acoust. Soc. Am. **48,** 950–965.

Yates, G. K., Winter, I. M., and Robertson, D. (**1990**). ''Basilar membrane non-linearity determines auditory rate-intensity functions and cochlear dynamic range,'' Hearing Res. **45,** 203–220.

Zagaeski, M., Cody, A. R., Russell, I. J., and Mountain, D. C. (**1994**). ''Transfer characteristic of inner hair cell synapses: Steady-state analysis,'' J. Acoust. Soc. Am. **95,** 3430–3434.

Zenner, H. P. (**1986**). ''Motile responses of outer hair cells,'' Hearing Res. **22,** 83–90.

Developmental changes in the precedence effect: Estimates of minimum audible angle

Ruth Y. Litovsky[a]

Psychology Department, University of Massachusetts, Amherst, Massachusetts 01003

The precedence effect refers to an auditory phenomenon which occurs when two similar sounds are presented from different locations with a brief delay, and only one sound is heard whose perceived location is dominated by the first source. Although the lagging source is not localized as an independent event, under some conditions, adults are able to extract its directional cues. Developmental studies suggest that this ability changes during development. However, those studies have used stimulus configurations which minimize the measurement of that ability. In the present study adults were first tested under several conditions, and the one which produced optimal performance was chosen for testing children. Using the minimum audible angle (MAA) task in the azimuthal plane, performance was compared for a single-source condition and two precedence conditions: in lag discrimination the lagging source changed location while the lead remained at midline, and in lead discrimination the reverse occurred. Subjects were 18 months old, 5 years old, and adult. Significant improvements in MAA occurred with an increase in age, especially in the precedence conditions. Within each group, performance was significantly better in single-source condition, followed by the lead and the lag discrimination. © *1997 Acoustical Society of America.* [S0001-4966(97)03709-0]

PACS numbers: 43.66.Ba, 43.66.Mk, 43.66.Pn [RHD]

INTRODUCTION

The precedence effect refers to an auditory phenomenon which occurs when two similar sounds are presented from different locations with a brief delay, and only one sound is heard whose perceived location is dominated by the first source (Wallach *et al.*, 1949; Blauert, 1983; Zurek, 1987). Although listeners are aware of the presence of the second source, under many conditions they find difficulty in extracting its directional information. While this difficulty is high when the leading and lagging sources are on opposite hemifields and separated by 80° or by large interaural-time differences (Wallach *et al.*, 1949; Zurek, 1980; Yost and Soderquist, 1984; Freyman *et al.*, 1991; Divenyi, 1992), it is reduced when the two sources are nearer (Perrott *et al.*, 1989; Saberi and Perrott, 1990; Shinn-Cunningham *et al.*, 1993; Litovsky and Macmillan, 1994; Litovsky *et al.*, 1996).

Developmental studies with humans suggest that the precedence effect is not present at birth, but appears at 4 to 5 months of age. At that age, the delay at which the lagging source is localized as an independent auditory event is longer than it is at 5 years or adult (Clifton, 1985). However, 5-year-olds' performance is only similar to adults' for simple, transient stimuli, such as clicks, but worse for longer, more complex stimuli (Morrongiello *et al.*, 1984). Hence 5 years of age may reflect a transitional stage in the development of the precedence effect (Clifton, 1985; Litovsky and Ashmead, 1997).

Interpretations of these findings are tricky, for the data suggest that when the precedence effect does appear during infancy it may be *stronger* than it is in children or adults, although increased thresholds may simply be due to the infants' inability to extract directional information from the lagging source. Note that in those studies the leading and lagging stimuli were presented from opposite hemifields, maximizing the physical separation between them. Since this stimulus configuration is thought to reduce listeners' ability to extract directional information from the lagging source, the present study was aimed at investigating conditions under which performance is optimal, at least in adults. The task of choice, which has been used extensively with single-source stimuli in young infants and children (e.g., Ashmead *et al.*, 1987; Morrongiello, 1988) was the minimum audible angle (MAA) in the azimuthal plane, estimating the smallest lateral difference in the position of a sound that can be detected reliably (Mills, 1958). Since the tasks used in previous developmental studies on precedence only required that listeners identify the hemifield containing the lag, more precise localization was not measured. Thus an additional benefit of the MAA task is that it allows one to measure developmental changes in localization precision under conditions in which the precedence effect occurs.

Adults were first tested on a classic *fusion* task, in which they reported whether they heard one fused auditory image or two separate sound sources, for lead-lag delays ranging from 2 to 12 ms. The duration of each burst was either 4 or 25 ms, and the longest delay at which adults reported hearing one source on less than 25% of trials was chosen for the MAA procedure. The data were then compared with those of Litovsky and Macmillan (1994) who used 6-ms noise bursts. Finally, the stimulus duration for testing children was chosen to match the one that resulted in the smallest MAAs in adults' lag-discrimination. Figure 1 illustrates the three con-

---

[a]Current address: Dept. of Biomedical Engineering, Boston University, 44 Cummington St., Boston, MA 02215, Electronic mail: Litovsky@enga.bu.edu
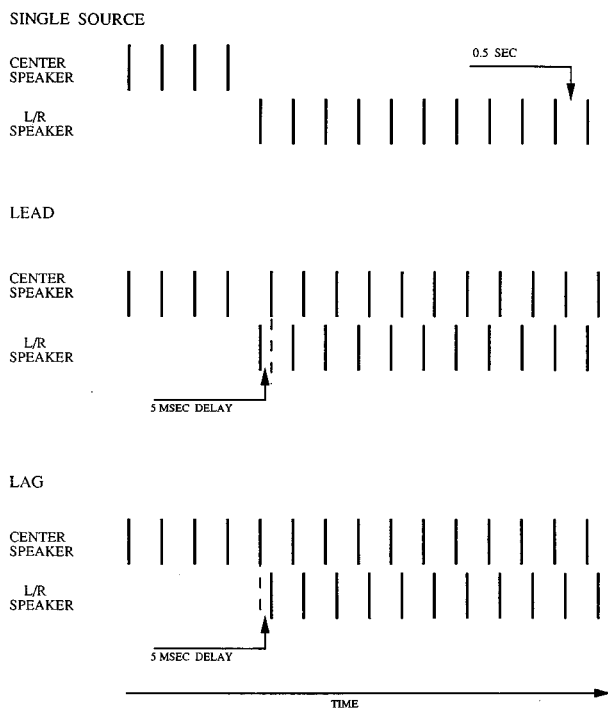
FIG. 1. Configuration of auditory stimuli. Three trial types are shown: single source, lead discrimination, and lag discrimination. All three trial types began with a single stimulus from the center loudspeaker, presented at a rate of 2/s. In single-source trials a single stimulus was then presented from either the right or left speaker. In lead and lag discrimination there were pairs of noise bursts with a 5-ms delay. In lead discrimination the leading source was presented from the left or right and the lagging source from the center. In lag discrimination the lagging source came from the left or right and the leading source from the center.

ditions that were used for the MAA tasks. In the single-source condition, a sound was presented from the middle of three speakers, and shifted randomly to one of two speakers (L or R). The other two conditions extended the task to precedence-effect stimuli: In lead discrimination, the leading source was presented to the left or right, and the lagging source from the middle. In lag discrimination, the lagging source was presented from either the left or right, and the leading source from the middle. While lag discrimination probes listeners' ability to discriminate changes in the location of the lagging source, lead discrimination reflects listeners' ability to overcome potentially confusing directional information from the lagging source.

Subjects were tested at ages 18 months, 5 years, and adult. At 18 months, MAAs are low (4°; Morrongiello, 1988) compared with adults (1° to 2°; e.g., Gardner, 1968; Mills, 1958; Hartmann and Rakerd, 1989; Saberi and Perrott, 1990; Litovsky and Macmillan, 1994). However, at 18 months, the precedence effect may not be fully developed due to incomplete maturation of the auditory cortex (Dekaban, 1970) which is thought to be involved in the ability to perform on some precedence tasks (Cranford and Oberholtzer, 1976; Whitfield et al., 1978; Clifton, 1985). Five years of age may reflect transition in the development of binaural hearing, hence it is ideal for measuring developmental changes in source discrimination under conditions of the precedence effect. In addition, MAAs with single-source stimuli have not

been previously measured at this age, hence this study bridges a gap in published data regarding changes in MAA between infancy and adulthood.

I. METHOD

A. Subjects and design

None of the children or adults had a known history of hearing impairment. All children were tested using the stimulus with 25-ms duration. Children groups included thirty-six 18-month-olds (12 males, 24 females; mean age of 18 months, 3 weeks), and thirty-six 5-year-olds (19 males, 17 females; mean age of 5 years, 4.4 months). Within each age subjects were randomly assigned to one of three groups ($N = 12$ each) according to stimulus type (single source, lead or lag discrimination).

Adult subjects were undergraduate students at the University of Massachusetts. Their hearing was screened so that pure-tone sensitivity matched in the two ears within 10 dB or less, with detection levels no more than 20 dB above those of normal levels at frequencies ranging between 250 and 8000 Hz. Twenty-four subjects (8 males and 16 females; mean age=20 yr, ranging between 19–21 yr) were tested, 12 in the 4-ms condition and 12 in the 25-ms condition. Each adult listener was tested on single source, lead and lag discrimination, with the order of the three stimulus conditions randomly assigned.

B. Auditory stimuli

Stimuli were 4- or 25-ms wideband (500 to 8500 Hz) noise bursts with 2-ms rise–fall times. Pilot testing in our lab as well as by others (Zurek, personal communication) suggests that the precedence effect may depend on the token of noise chosen. Hence on each trial the noise bursts were selected independently from a long segment of the noise. In lead and lag discrimination trials the two bursts consisted of the same token of noise. Stimuli were computer generated with 16-bit precision, converted to analog form at 20 kHz (TTES-QDA1), low-pass filtered at 8500 Hz (TTE J1390), and tape-recorded (Teac X-300). During testing the prerecorded stimuli were amplified and played back from the same tape recorder over loudspeakers. The sounds were presented at A-weighted levels of 50–52 dB (B&K 2204 SLM) over a background level of 28 dBA, as measured at the approximate position of the subject's head. The time sequence for a trial for each of the three stimulus conditions is presented in Fig. 1. Each trial consisted of 15 noise bursts, presented at a rate of 2/s. In the single-source condition the first four noise bursts were presented from midline, followed by 11 noise bursts from either the right or left speaker. Lead- and lag-discrimination trials also began with four single-source noise bursts from midline. In the 11 bursts that followed there were two noise samples per burst, with the onset of one delayed relative to the onset of the other by 5 ms. In lead discrimination the leading source came from the right or left and the lagging source from the middle; in lag discrimination the opposite occurred.

C. Apparatus

The study was conducted in a sound-attenuated room, 3.5×4.0 m. Subjects sat facing an arc-shaped apparatus spanning 110° of an imaginary circle in the azimuthal plane with the subject at the center of a 1.65-m radius. Adults and 5-year-olds sat on a chair and 18-month-olds were seated on their parent's lap. Parents wore masking head phones which obscured all directional information to avoid the possibility of them systematically cueing their children in either direction. The arc was covered by a dark curtain to hide the three loudspeakers (Radio Shack model Minimus-7). Speaker responses were measured with a sound-level meter (B&K) at the approximate position of the listener's head, and all speakers had matching frequency responses within 1 to 2 dB for all frequencies between 100–8000 Hz. During all trials one loudspeaker remained at midline while the other two were positioned at equal angles to the left and right of midline.

Adjoining the testing room was a control room from which the session was monitored. A video deck (Panasonic GX2 1950) and television monitor in this room received input from a video camera inside the testing chamber. The investigator in this room viewed the subject's behavior and administered reinforcement following correct responses. Two identical sets of reinforcers were positioned at 60° to the left and right; each set consists of two mechanical toys which, when activated, provided a visual/auditory display known to be attractive to infants (Trehub *et al.*, 1981). Each toy was enclosed within a smoked-plexiglass box so that it remained invisible to the subjects except when activated. A video camera was positioned above the curtain at midline with output to one monitor behind the curtain, and another monitor in the outside control room. This double output allowed both experimenters to view the subjects' behavior during the session, which was especially important for testing 18-month-olds, whose responses were measured in terms of correct head-turning behaviors towards the appropriate loudspeaker.

D. Procedure

1. Testing adults in the fusion experiment

In order to establish which delays were most appropriate for measuring MAAs using lead-lag noise pairs, adults' perception of whether the lead and lag were fused was measured. On each trial the lead and lag were presented from 0° (front) and 30° right, respectively. This source separation of 30° was chosen so that it exceeded the MAAs of all age groups. The delays between the lead and lag included 2, 4, 6, 8, 10, and 12 ms. Within each block of trials there were 60 trials, consisting of ten repetitions of each delay, presented in random order. Each block was repeated five times, for a total of 50 trials per delay. On each trial listeners were instructed to report whether they perceived one fused auditory image, or two separate sound sources. The longest delay at which adults reported hearing one source on less than 25% of trials was chosen for testing on the MAA procedure. The aim was

to identify a delay that is below echo threshold (Blauert, 1983), at which there might be some effect of both the lead and lag on sound localization precision.

2. Testing procedure for 18-month-olds on the MAA task

Experimenters wore masking earphones to avoid cueing subjects regarding stimulus location. Each trial was initiated by attracting the subject's attention to the center speaker. The stimulus was initiated once the child was facing the center speaker directly. An experimenter ''judged'' whether the subject's behavior indicated a shift of the sound to the right or left. A correct judgment, and hence response, resulted in activation of a reinforcer on the correct side for 5 s. An incorrect judgment resulted in a time-out period of 5 s. If no head turn was made during the 5 s after the stimulus shifted from midline, the trial was considered a nonresponse trial in which no reinforcement was delivered and a 5 s time-out period ensued. Three observers were trained on judging head turning responses. Percent agreement for the three observers for all judgments in this study were: A and $B = 95\%$; A and $C = 96\%$; B and $C = 94\%$.

3. Testing procedure for 5-year-olds and adults on the MAA task

Subjects were asked to center their heads and look at a target positioned at midline at the onset of each trial. They were instructed to point toward the right or left hemifield once the sound shifted away from midline, and to guess if they did not perceive an obvious change. Following a correct response children were presented with the toy that was used with 18-month-olds on the correct side. Adults were given feedback concerning the correct side by activation of a light bulb. For both ages incorrect responses were followed by a 5-s time-out period and no other feedback. Prior to test trials, subjects had to meet the criterion of correct responses on four out of five consecutive single-source practice trials with loudspeakers at 55°, and were allowed a maximum of ten trials to reach criterion. Ten children were excluded from the final sample due to suspicion of hearing impairment ($N = 2$) or loss of interest in the task ($N = 8$). Four adult subjects were tested but excluded from the final sample due to failure on the screening hearing test.

4. Adaptive method and MAA estimation

Changes in angular separation of the loudspeakers were determined using the classic two-down/one-up method of Levitt (1971) which seeks the 71% correct point on a psychometric function. The initial angles were chosen to yield high accuracy at each age-stimulus combination, as determined during pilot testing. For single source and lead discrimination with 18-month-olds and all stimuli with 5-year-olds, the starting angle was 55°. For lag discrimination with 18-month-olds it was 75°, and for adults in all conditions it was 30°. Step sizes of angular change were determined by a modified version of PEST (Macmillan and Creelman, 1991, see Chap. 8; Litovsky and Macmillan, 1994), with the following additional rules for increased estimation accuracy
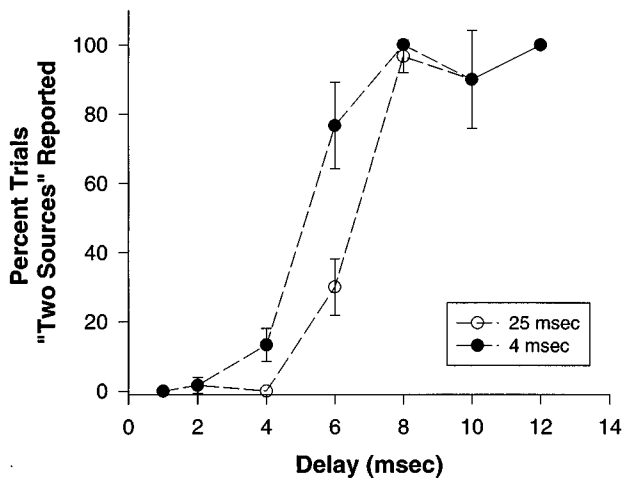
FIG. 2. Percent of trials on which listeners reported hearing "two sources" are plotted as a function of delay, for noise bursts that are either 25 or 4 ms in duration. Lead and lag sources were at 0° (front) and 30° right, respectively. Data are average and standard deviations for three listeners.

with children: (1) After two consecutive failures or nonresponse trials, a "probe" trial (Aslin *et al.*, 1981) was presented with the loudspeakers placed at the initial angle position. This trial type was repeated until a correct response was made, but data were not included in the estimation of MAAs. Once a correct response was made testing resumed at the angle position of the last failure. (2) A nonresponse trial was repeated at the same angle as the last and was not considered in MAA estimation. Testing was terminated once seven reversals were reached. The mean number of trials required to estimate MAA thresholds were 28.7 (range=14–50) for 18-month-olds, 27 (range=21–40) for 5-year-olds, and 26.5 (range=19–36) for adults. The proportions of nonresponse trials were 0% for adults, 1% for 5-year-olds, and 15% for 18-month-olds. MAA was estimated using maximum-likelihood rules that have been used extensively in combination with PEST (Macmillan and Creelman, 1991).

II. RESULTS

Shown in Fig. 2 is the percent of trials on which adults reported hearing two sources (rather than one fused auditory event), plotted as a function of delay for conditions in which the lead and lag noise bursts were either 4 ms in duration (filled circles) or 25 ms in duration (open circles). Listeners perceived the lagging source as an independent sound source on more than 50% of trials at a delay of 6 ms for the shorter-duration stimulus, and a delay of 8 ms for the longer-duration stimulus.

MAAs were averaged over subjects at each age for each stimulus condition. Where stated, statistical comparisons were conducted with *t* tests; significance values were set to 0.01 after applying Scheffe's adjustment for *post-hoc* contrasts. The means and standard deviations for adults are plotted in Fig. 3, comparing results in the 4- and 25-ms stimulus conditions. An additional set of data are replotted from Litovsky and Macmillan (1994) who used a 6-ms stimulus and tested subjects under identical conditions in the same room with the same apparatus. Stimulus duration has no significant
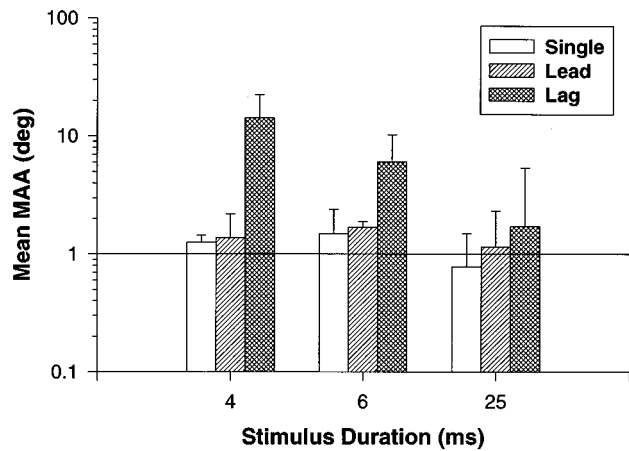
FIG. 3. Mean MAA estimates for adults, comparing single source, lead and lag discrimination, using 4-, 6-, and 25-ms stimuli. Data for the 6-ms condition are replotted from Litovsky and Macmillan (1994).

effect on MAAs for single source and lead discrimination. In contrast, lag-discrimination MAAs decrease significantly with longer durations.

Plotted in Fig. 4(A) are data from the three age groups for the 25-ms stimulus, comparing single-source, lead and lag-discrimination conditions (adult data from Fig. 3 are re-
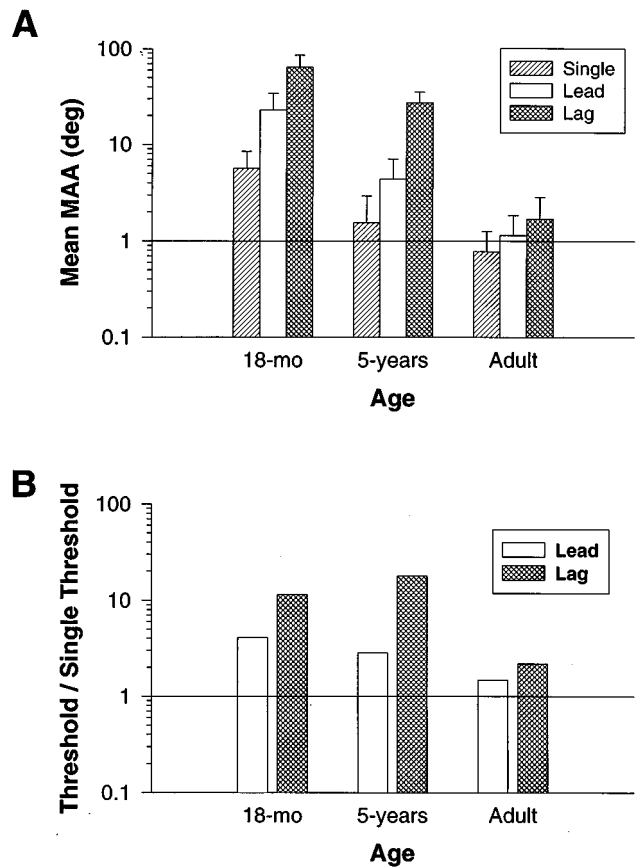
FIG. 4. (A) Mean MAA estimates for the 25-ms stimulus are plotted for each age group comparing single source, lead and lag discrimination. (B) Data from (A) were normalized within each age group by the mean MAA values obtained in the single-source condition.

plotted). Average MAAs (and standard deviations) for adults, 5-year-olds, and 18-month-olds, respectively, were 0.78 (0.48), 1.55 (1.38), and 5.65 (2.83) for single source; 1.15 (0.7), 4.40 (2.7), and 23.05 (11.07) for lead discrimination; and 1.71 (1.16), 27.5 (8.02), and 64.58 (21.6) for lag discrimination. Two aspects of the data are most noticeable. First, there is a general decrease in MAAs with an increase in age. Second, for both 18-month-olds and 5-year-olds, MAAs are substantially higher in the precedence conditions than in the single-source condition, with lag discrimination being especially high. Figure 4(B) shows the same data, normalized at each age by the mean MAA in the single-source condition. Developmental differences are still robust, maintaining the same trend that was observed in Fig. 4(A).

Statistical analyses showed that adults' MAAs are significantly lower in single source discrimination than lead discrimination ($t=2.9$), and than lag discrimination ($t=3.7$). MAAs are also lower in lead compared with lag discrimination ($t=3.3$). Thus compared with a single stimulus situation, adults' discrimination is degraded under conditions of the precedence effect, regardless whether the lead or lag is being localized. However, adults are better able to extract directional information from the leading source than from the lagging source.

Five-year-olds and 18-month-olds performed significantly better (i.e., lower MAAs) in the single source condition than either lead ($t=3.4$ and $t=6.3$, respectively) or lag discrimination ($t=3.3$ and $t=9.2$, respectively). Finally, their performance was better in lead than in lag discrimination ($t=2.9$ and $t=5.7$, respectively). Comparisons between the two age groups showed that 5-year-olds are significantly better than 18-month-olds in single source ($t=4.5$), lead discrimination ($t=5.7$), and lag discrimination ($t=3.6$), suggesting that between the ages of 18 months and 5 years children's localization precision improves significantly. Adults perform significantly better than 18-month-olds in single source ($t=5.9$), lead discrimination ($t=6.8$), and lag discrimination ($t=10.0$). They also perform better than 5-year-olds on lead and lag discrimination ($t=4.0$ and $t=3.2$, respectively). However, 5-year-olds do not perform significantly worse than adults in the single-source condition, implying that basic localization precision may have reached adult acuity by childhood, whereas precision under conditions of the precedence effect has not.

III. DISCUSSION

A. MAAs with single-source stimuli

Adult MAAs (mean=0.78°) are consistent with previous findings for broadband stimuli (Perrott *et al.*, 1989; Saberi and Perrott, 1990). The estimate of 5.7° for 18-month-olds is slightly higher than a previous reported value of 4.0° (Morrongiello, 1988). The difference of 41% between Morrongiello's (1988) results and the present study cannot be attributed to differences in target proportions since both studies used approximately 71% correct to estimate thresholds. A major difference however, is Morrongiello's use of the ''method-of-constant-stimuli,'' compared with the adaptive method used here. In fact, MAA estimates in 6-month-old

infants have also been higher with the adaptive method (19°; Ashmead *et al.*, 1987) than the method of constant stimuli (12°; Morrongiello, 1988).

The most noteworthy finding is that 5-year-olds, on average, do not perform significantly worse than adults on the single-source MAA task. Despite the fact that 5-year-olds' mean MAA (1.55°) was twofold that of adults (0.78°), the difference was not statistically significant. The mean values are within the range of previously reported MAAs in adults (e.g., Gardner, 1968; Hartmann and Rakerd, 1989; Perrott *et al.*, 1989; Litovsky and Macmillan, 1994). It may be interesting to note that five of the 12 children actually had MAAs below 1°, while two childrens' MAAs were above 4°, which fall closer to the mean of the 18-month-old group. Hence, five years of age may represent a transitional stage during which some children have reached adult acuity in the discrimination task while other children have not. Alternatively, the variability within the 5-year-olds may represent individual differences in attentional capacities.

Nonetheless, the author is not aware of previous MAA estimates for children at any age above 24 months (Morrongiello, 1988). The value of 1.55° does fall, as expected, between MAAs of 18-month-olds and adults, suggesting that development in localization precision continues to occur between the second and fifth years of life. Since previous reports suggest that by this age children have not yet reached adult-level performance on other tasks involving temporal discrimination (Davis and McCroskey, 1980; Irwin *et al.*, 1985; Wightman *et al.*, 1989), localization precision for single-source stimuli may not depend solely on temporal acuity.

B. MAAs under conditions of the precedence effect

In the adult group the differences in means between single source, lead and lag discrimination were quite small, however, all three conditions were significantly different from one another. The finding that lead-discrimination MAAs were worse than single-source MAAs suggests that the presence of the lagging source at midline did interfere somewhat with listeners' ability to extract directional information from the lead. Thus precedence in this situation was not ''perfect.'' However, precedence did exist to the extent that lag discrimination was worse than lead discrimination. These results are consistent with previous findings on MAAs under conditions of the precedence effect (Perrott *et al.*, 1989; Litovsky and Macmillan, 1994). It has long been suggested that the precedence effect is an auditory phenomenon that diminishes the influence of directional information from echoes, thereby aiding an organism in accurately localizing the original sound source (Zurek, 1980, 1987). The finding that performance was, at all ages, significantly better in lead than lag discrimination is consistent with this notion, and with previous reports that the precedence effect gives dominance to directional cues provided by the leading source. One such measure is a just-noticeable difference in the interaural-time difference of a signal, which is not affected when the signal is the leading source, but is strongly affected when it is the lagging one (e.g., Wallach *et al.*, 1949; Zurek, 1980; Yost and Soderquist, 1984; Shinn-Cunningham *et al.*,

1993). A tentative interpretation of these results is a masking of the ''localization strength'' of the lagging source (Divenyi, 1992).

1. Effect of stimulus duration in adults

The finding that longer-duration stimuli result in increased fusion, i.e., that listeners require longer temporal separation between the lead and lag before the lag is heard, is consistent with previous reports in the literature. For example, echo thresholds for click stimuli (2–5 ms) are lower than those for noise stimuli (>8 ms; see Schubert and Wernick, 1969 and Blauert, 1983). Recent physiological studies showing neural correlates of the precedence effect in the brainstem of the cat (Yin, 1994; Fitzpatrick et al., 1995; Litovsky et al., 1997; Litovsky and Yin, 1993, 1997) have also found that single neurons' echo thresholds increase with longer-duration noise stimuli. The physiological findings suggest that, although both the leading and lagging responses produce increased excitability when the stimulus duration is longer, the end result is stronger suppression of the lagging response. Thus changes in duration seem to have a more potent effect on the amount of suppression produced by the lead than on the excitation produced by the lag.

The finding that longer-duration stimuli result in decreased lag-discrimination MAAs may not be a function of mechanisms that are involved in precedence per se. To date, there are no published data known to the author with which to compare these findings. A possible explanation is that the longer-duration stimuli provide a temporal ''tail'' consisting of the lagging source; the leading stimulus has been turned off but the lagging one continues for 5 ms. Adults may have developed a listening strategy that enables them to extract enough directional information from the tail-end of the stimulus. However, that strategy may have only succeeded in the 25-ms condition when the lagging source, after having overlapped with the lead for 20 ms, was presented by itself for 5 ms. In the lead-discrimination task listeners only had to pay attention to the beginning of the stimulus, regardless of what came after its onset, hence there were no differences in the MAAs for the different stimulus durations.

2. Developmental changes in the precedence effect

Five-year-olds performed significantly better than 18-month-olds, and adults performed better than both children groups in the two precedence tasks. The finding that developmental changes remained fairly constant after normalizing lead- and lag-discrimination MAAs by single-source MAAs [Fig. 4(B)], further suggests that MAAs under conditions of the precedence effect are not merely a by-product of a ''noisy'' single-source discrimination system.

The lag-discrimination data suggest that with an increase in age there is an improvement in listeners' ability to extract directional information from a simulated echo, i.e., a sound source that is not localized as an independent auditory event (lag discrimination). This finding is consistent with the developmental data of Clifton and colleagues (see Clifton, 1985) who measured listeners' ability to localize the hemifield containing the lagging source, and found elevated thresholds in infants and children compared with adults. That is, infants and children required longer delays between the lead and lag before they could localize the lag as a separate source, at its respective location. The higher fusion point for younger listeners could be the primary cause for the difficulty they exhibited in the lag-discrimination task.

These results might reflect different maturational stages in the auditory cortex (Clifton, 1985), which has been implicated as a necessary structure for the binaural suppression of echoes (Cranford and Oberholtzer, 1976; Whitfield et al., 1978). The brain undergoes substantial growth during early childhood, with the most dramatic changes occurring in cortical regions (Dekaban, 1970; Yakovlev and Lecours, 1967). Although differences in attentional or learning processes cannot be ruled out, the nature of the precedence effect would seem to favor the cortical development explanation. Precedence is a very compelling auditory illusion, and in everyday listening situations adults seem capable of suppressing echoes in a natural manner, regardless of whether they have cognitive knowledge of the presence of those echoes. However, we can be trained to ''hear out'' the echoes and extract information from them (e.g., Saberi and Perrott, 1990). Do these findings suggest that adults have ''weaker'' precedence than children? Perhaps it is simply the case that adults are better and more experienced than children at most tasks, including ones which involve attending to directional changes in a simulated echo.

The lead-discrimination data suggest that with an increase in age listeners are better able to ignore directional information from the lagging source when that information is irrelevant to the task. Under conditions of the precedence effect adult listeners are almost always aware of the presence of the lagging source and they can discriminate between a single-source sound and a paired-source sound (Blauert, 1983; Zurek, 1987). However, the potency of precedence lies in the fact that the auditory system fails to assign a separate location to the lagging source. This ability, too, might depend on the development of sensory and cortical structures in the auditory system. Alternatively, it might rely on the attainment of cognitive and attentional skills, which undergo significant changes during early childhood (Werner, 1992; Litovsky and Ashmead, 1997).

Finally, it is interesting that 5-year-olds' single-source MAAs have reached adult-level maturity, whereas their MAAs with paired sounds have not. Possibly, this reflects a decoupling between the development of basic auditory abilities required for single-source discrimination and perhaps more sophisticated skills such as accommodating echoes. It is also possible that localization precision requires special cognitive skills in the presence of echoes that may not be available to young children. An alternative explanation may relate to stimulus parameters. Five-year-olds' ability to identify the hemifield containing the lag is only similar to adults' for click stimuli; it is significantly different than adults' for long-duration noise (Morrongiello et al., 1984). The present study only tested children with the long-duration stimuli; it is possible that with short-duration stimuli performance on lead discrimination might have been similar to adults.

C. Summary

In summary, it is clear that several aspects of localization precision are undergoing developmental changes during infancy and childhood. These changes are useful for understanding the relationship between two aspects of the precedence effect, fusion, and suppression of directional cues. This study found a discrepancy between the ages at which precision matures for simple and complex tasks, suggesting perhaps, that different mechanisms might be responsible for the development of localizing single source and paired stimuli. Finally, it must be acknowledged that nonsensory factors related to the task may have contributed to developmental differences. It is therefore unlikely that precedence is merely a by-product of simple sound localization. A true test of this issue may require studies on the neurophysiological basis of precedence.

Ashmead, D. H., Clifton, R. K., and Perris, E. E. (**1987**). ''Precision of auditory localization in human infants, '' Dev. Psych. **23**, 641–647.

Aslin, R. N., Pisoni, D. B., Hennesy, B. L., and Perey, A. J. (**1981**). ''Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience,'' Child Dev. **52**, 1135–1145.

Blauert, J. (**1983**). *Spatial Hearing* (MIT, Cambridge, MA).

Clifton, R. K. (**1985**). ''The precedence effect: Its implications for developmental questions,'' in *Auditory Development in Infancy*, edited by S. E. Trehub and B. Schneider (Plenum, New York).

Cranford, J. L., and Oberholtzer, M. (**1976**). ''Role of Neocortex in binaural hearing in the cat, II: The precedence effect in sound localization,'' Brain Res. **111**, 225–239.

Davis, S., and McCroskey, R. (**1980**). ''Auditory fusion in children,'' Child Dev. **51**, 75–80.

Dekaban, A. (**1970**). *Neurology of Early Childhood* (William and Wilkins, Baltimore).

Divenyi, P. L. (**1992**). ''Binaural suppression of nonechoes,'' J. Acoust. Soc. Am. **91**, 1078–1084.

Fitzpatrick, D. C., Kuwada, S., Batra, R., and Trahiotis, C. (**1995**). ''Neural responses to simple, simulated echoes in the auditory brainstem of the unanesthetized rabbit,'' J. Neurophysiol. **74**, 2469–2486.

Freyman, R. L., Clifton, R. K., and Litovsky, R. L. (**1991**). ''Dynamics of the precedence effect,'' J. Acoust. Soc. Am. **90**, 874–884.

Gardner, M. B. (**1968**). ''Historical background of the Haas and/or precedence effect,'' J. Acoust. Soc. Am. **43**, 1243–1248.

Hartmann, W. M., and Rakerd, B. (**1989**). ''On the minimum audible angle-A decision theory approach,'' J. Acoust. Soc. Am. **85**, 2031–2041.

Irwin, R. J., Ball, A. K. R., Kay, N., Stillman, J. A., and Rosser, J. (**1985**). ''The development of auditory temporal acuity in children,'' Child Dev. **56**, 614–620.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Litovsky, R. Y., and Ashmead, D. M. (**1997**). ''Developmental aspects of binaural and spatial hearing,'' in *Binaural and Spatial Hearing*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum, Hillsdale, NJ).

Litovsky, R. Y., and Macmillan, N. A. (**1994**). ''Minimum audible angle for clicks with simulated echoes: Effects of azimuth and standard,'' J. Acoust. Soc. Am. **96**, 752–758.

Litovsky, R. Y., and Yin, T. C. T. (**1993**). ''Single-unit responses to stimuli that mimic the precedence effect in the inferior colliculus of the cat,'' Assoc. Res. Otolaryng. (Abstract).

Litovsky, R. Y., and Yin, T. C. T. (**1997**). ''Physiological studies of the precedence effect in the inferior colliculus of the cat: I. Correlates of psychophysics,'' J. Neurophysiol. (submitted).

Litovsky, R. Y., Dizon, R., Pazmany, C., and Colburn, H. S. (**1996**). ''Studies of the precedence effect in the median sagittal and azimuthal planes in a virtual acoustic space,'' Assoc. Res. Otolaryng. (Abstract).

Litovsky, R. Y., Rakerd, B., Yin, T. C. T., and Hartmann, W. M. (**1997**). ''Evidence for psychophysical and physiological correlates of the precedence effect in the sagittal plane,'' J. Neurophysiol. **77**, 2223–2226.

Macmillan, N. A., and Creelman, C. D. (**1991**). *Detection Theory: A User's Guide* (Cambridge U.P., Cambridge, England).

Mills, A. (**1958**). ''On the minimum audible angle,'' J. Acoust. Soc. Am. **30**, 237–246.

Morrongiello, B. (**1988**). ''Infants' localization of sounds along the horizontal axis: Estimates of minimum audible angle,'' Dev. Psych. **24**, 8–13.

Morrongiello, B., Kulig, J., and Clifton, R. (**1984**). ''Developmental changes in auditory temporal perception,'' Child Dev. **55**, 461–471.

Perrott, D. R., Marlborough, K., Merrill, P., and Strybel, T. Z. (**1989**). ''Minimum audible angle thresholds obtained under conditions in which the precedence effect is assumed to operate,'' J. Acoust. Soc. Am. **85**, 282–288.

Saberi, K., and Perrott, D. R. (**1990**). ''Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate,'' J. Acoust. Soc. Am. **87**, 1732–1737.

Schubert, E. D., and Wernick, J. (**1969**). ''Envelope versus microstructure in the fusion of dichotic signals,'' J. Acoust. Soc. Am. **45**, 1525–1531.

Shinn-Cunningham, B. G., Zurek, P. M., and Durlach, N. I. (**1993**). ''Adjustment and discrimination measurements of the precedence effect,'' J. Acoust. Soc. Am. **93**, 2923–2932.

Trehub, S. E., Schneider, B. A., and Bull, D. (**1981**). ''Effect of reinforcement on infants' performance in an auditory detection task,'' Dev. Psych. **17**, 872–877.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). ''The precedence effect in sound localization,'' J. Am. Psychol. **57**, 315–336.

Werner, L. A. (**1992**). ''Interpreting developmental psychoacoustics,'' in *Developmental Psychoacoustics*, edited by L. A. Werner and E. W. Rubel (APA, Washington, DC).

Whitfield, I. C., Diamond, E., Chiverallis, K., and Williamson, T. (**1978**). ''Some further observations on the effects of unilateral cortical ablation on sound localization in the cat,'' Exp. Brain Res. **31**, 221–234.

Wightman, F., Allen, P., Dolan, T., Kistler, D., and Jamieson, D. (**1989**). ''Temporal resolution in children,'' Child Dev. **60**, 611–624.

Yakovlev, P., and Lecours, A. (**1967**). ''The myelogenetic cycles of regional maturation of the brain,'' in *Regional Development of the Brain in Early Life*, edited by A. Minkowski (Davis, Philadelphia).

Yin, T. C. T. (**1994**). ''Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat,'' J. Neurosci. **14**, 5170–5186.

Yost, W. A., and Soderquist, D. R. (**1984**). ''The precedence effect: Revisited,'' J. Acoust. Soc. Am. **76**, 1377–1383.

Zurek, P. M. (**1980**). ''The precedence effect and its possible role in the avoidance of interaural ambiguities,'' J. Acoust. Soc. Am. **67**, 952–964.

Zurek, P. M. (**1987**). ''The precedence effect,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York).

Zurek, P. M. (personal communication). ''Findings on precedence and 'anti-precedence' that vary with different noise tokens.''

Signal detection and pitch ranking in conditions of masking release

Joseph W. Hall, John H. Grose, and Madhu B. Dev
*The Division of Otolaryngology/Head & Neck Surgery, University of North Carolina Medical School,
Chapel Hill, North Carolina 27599-7070*

Masked threshold levels for signal detection or pitch ranking (low, middle, or high) were determined in conditions where one of three possible signal frequencies could be presented on a given trial of a three-interval forced-choice task. Thresholds were determined under conditions investigating binaural masking release (the masking-level difference, or MLD) and monaural masking release in modulated noise. It was assumed that part of the masking release in modulated noise was due to a within-channel analysis of information in the masker dips, and that part of the masking release was due to across-frequency analysis of temporal envelope information (comodulation masking release, or CMR). In the MLD experiment, the masker was a low-pass noise with a frequency cutoff of 1000 Hz. In the modulated-noise experiment, a broadband noise was square-wave amplitude modulated at a rate of 25 Hz and then digitally low-passed filtered at 1000 Hz. The results indicated that when the frequency separation between frequency components was relatively wide, masking release for pitch ranking was similar to that for signal detection. However, as the frequency separation between components narrowed, masking release for pitch ranking decreased. The results are consistent with an interpretation that information regarding signal frequency is relatively coarse under the conditions of masking release examined here. © *1997 Acoustical Society of America.*
[S0001-4966(97)03009-9]

PACS numbers: 43.66.Dc, 43.66.Fe, 43.66.Hg, 43.66.Pn [WJ]

INTRODUCTION

The present study investigated masking release in which one of three possible signal frequencies was presented during one interval of a three-interval forced-choice procedure. The task of the listener was either to detect the signal, or to rank its pitch (as low, middle, or high). Masking release was examined both for a binaural paradigm using the masking-level difference (MLD) (Hirsh, 1948; Jeffress *et al.*, 1952), and a monaural paradigm comparing performance in unmodulated and modulated masking noise (Zwicker and Schorn, 1982; Hall *et al.*, 1984; Carlyon *et al.*, 1989; Humes, 1990). It has been hypothesized that part of the reason that detection thresholds are better in modulated than in unmodulated noise, when the noise bandwidth is relatively wide, is related to comodulation masking release, or CMR (Hall *et al.*, 1984; Carlyon *et al.*, 1989). In this interpretation, the masker dips as processed by auditory filters remote from the signal frequency may cue the best listening epochs for processing signal energy (Buus, 1985).

Whereas most studies of MLD and CMR have examined signal detection, only a small number of studies have examined how these types of masking release pertain to the processing of suprathreshold stimuli. There is evidence that both of these types of masking release are greater when measured for detection threshold than for performance that requires the analysis of suprathreshold signals. For example, both MLD and CMR have been shown to be relatively small in amplitude discrimination and temporal gap detection paradigms (Henning, 1973, 1991b, 1991a; Hall and Grose, 1992, 1995). MLD has also been shown to be small in tasks investigating frequency discrimination (Gebhardt *et al.*, 1971; Henning,

1991b) and speech intelligibility (Levitt and Rabiner, 1967). Likewise, CMR was found to be relatively small in a vowel identification task, and absent in a more difficult speech recognition task involving the recognition of monosyllabic words (Grose and Hall, 1992). Hall and Grose (1995) suggested that whereas the mechanisms accounting for CMR provide information sufficient for detection at low signal-to-noise ratios, the information is in some sense coarse or impoverished in relation to the information available in the nonmasking release condition at a comparable sensation level (SL).

The present study examined MLD and masking release in modulated noise under conditions where the task of the listener was either to detect a pure-tone signal, or to rank its pitch. In conditions examining MLD, the masking noise was presented interaurally in phase to the two ears (No), and the signal was presented either in phase (So) or 180° out-of-phase (Sπ). MLD is defined as the threshold for the NoSo condition minus the threshold for the NoSπ condition. In conditions involving modulated noise, the noise and signal were presented monaurally, and the noise was either unmodulated or was amplitude modulated. For both detection and pitch ranking, the signal was one of three possible frequency components presented during one of three intervals on a given trial. In signal detection conditions, the task of the listener was simply to detect the presence of the signal. In pitch-ranking conditions, the task of the listener was to rank the pitch of the signal (low, middle, or high). Presumably, the task of pitch ranking would require better stimulus information than the simpler task of detection, as the listener would not only have to hear the signal, but also make a

judgment about its pitch. The frequency spacing between the three signals was an independent variable. The question of interest was how the magnitude of masking release would vary between the detection and ranking tasks, as a function of the frequency spacing between the signal frequencies. If stimulus information were relatively poor under conditions of masking release, then the amount of masking release might depend heavily on the frequency separation between the signals. If the frequency separation were large, then relatively poor information available near the detection threshold in modulated noise might be sufficient to perform the ranking task, and the masking release for pitch ranking would therefore be large. However, if frequency separation were small, the relatively poor information available near the detection threshold in modulated noise may not be sufficient to perform the ranking task, therefore requiring a higher signal-to-noise ratio.

For conditions investigating modulated noise, it was assumed that masking release could be due to both within-channel and across-channel analysis of envelope dip information. In signal detection conditions, the magnitude of the within-channel masking release can be estimated from thresholds in modulated and unmodulated narrow-band noise having a bandwidth similar to the width of the auditory filter at the signal frequency (Carlyon *et al.*, 1989). The masking release obtained for this narrow-band condition can then be subtracted from the masking release found for the wide-band condition, with the remainder taken as the approximate CMR (across-channel masking release). A corresponding procedure is not available for estimating the within-channel contributions to masking release for pitch ranking, simply because the paradigm of multiple signal frequencies requires a noise band that is wider than the auditory filter at any one of the set of three signal frequencies. In the present experiment, therefore, we did not specifically separate the within-channel and across-channel factors contributing to the masking release obtained in modulated noise. We did, however collect data to obtain an approximation of the within-channel masking release for signal detection.

In summary, the present experiment examined masking release for pitch ranking in both an MLD paradigm and a modulated-noise paradigm. Poor pitch ranking for binaural masking release can be expected because of previous data indicating relatively poor frequency discrimination in binaural masking release. Expectations for the modulated-noise paradigm are less clear. Although previous data indicated relatively small CMR for other suprathreshold measures (amplitude discrimination and gap detection), there are no previous data that we are aware of on frequency discrimination or pitch ranking in modulated noise. Pitch-ranking data on both binaural masking release and masking release in modulated noise were obtained so that results could be compared in similar conditions and listeners across these two types of masking release.

I. METHOD

A. Listeners

Five listeners with normal hearing (thresholds better than 20 dB HL for octave frequencies from 250 to 8000 Hz)

participated. Three of the listeners were male and two were female. Listener age ranged between 23 and 28 years. Each listener received approximately 4 h of training before data were collected on experimental conditions.

B. Procedure

Three-interval, three-alternative (3AFC) procedures were used to measure thresholds in both signal detection and pitch-ranking tasks. In the conditions measuring signal detection, one of three signal frequencies was presented in one interval of the 3AFC procedure, and the listener was asked to select the interval in which the signal occurred. The listener responded by pressing one of three buttons on a response box. In the conditions investigating pitch ranking, one of three signal frequencies was presented in one interval of the 3AFC procedure, and the listener was asked to rank the pitch (low, middle, or high) of the signal. The listener again responded by pressing one of three buttons on a response box. In all conditions, there was uncertainty with respect to both signal frequency and target interval.[1]

Adaptive threshold tracking methods (Levitt, 1971) estimating the 79.4% correct point on the psychometric function were used in both the detection and ranking tasks. Data were first obtained for the case of signal detection. In these conditions, three adaptive staircases were run simultaneously, one for each of the three signal frequencies. For each staircase, the attenuation value was decreased following an incorrect response, and was increased following three consecutive correct responses for that frequency. On any given trial, one of the three frequencies, at random, was presented. For each staircase, the initial stepsize of 8 dB was reduced to 4 dB after two reversals, and was reduced to 2 dB after two further reversals. All of the three staircases proceeded until at least ten reversals had occurred for each staircase, and the final six reversals of each staircase were averaged to estimate the threshold for a given signal frequency. Where one staircase reached its terminating criterion before another, presentation of the signal frequency associated with the "terminated" staircase continued to be included in the randomized presentation scheme, but the responses to the signal were disregarded. The final threshold for a condition was estimated as the average of four of these multiple staircase runs.

In the pitch-ranking conditions, only one staircase was used. On any given trial, any one of the three frequencies was presented. Again, attenuation was decreased after a single incorrect response, and was increased following three consecutive correct responses (note that three such consecutive responses could occur as a result of any combination of the three possible frequencies). For each listener, the relative intensity levels of the three frequency components in the pitch-ranking conditions was determined by the thresholds that had been obtained previously in the detection threshold conditions. For example, if masked detection thresholds of 50, 51, and 52 dB had been obtained for frequencies 1, 2, and 3, respectively, the starting levels for the pitch-ranking condition would be such that frequency 3 would have 2 dB less attenuation than frequency 1 and 1 dB less attenuation than frequency 2. Furthermore, this intensity relation was

maintained throughout the threshold run. This procedure was intended to insure that the audibility of the three components would remain roughly similar throughout the run, and that the listener's response would be based on the pitch of the stimulus rather than upon its audibility.[2] Again, the initial stepsize of 8 dB was reduced to 4 dB after two reversals, and was reduced to 2 dB after two further reversals. The procedure was terminated following 12 reversals, with the pitch-ranking threshold estimated as the average of the final 8 reversals. The final threshold for a condition was estimated as the average of the threshold estimates from four staircase runs. Pilot data (obtained on different subjects than those participating in the main experiment) did not indicate a bias for choosing a particular rank.

Stimuli were presented over Sony MDR V6 earphones. In the MLD conditions, stimuli were presented to both earphones, and in the modulated-noise conditions, stimuli were presented to the left ear. Listeners were seated in a double-walled Industrial Acoustics Corporation sound booth. Feedback was provided following each response.

C. Stimuli and conditions

There were five different frequency separation conditions, each involving three signal frequencies geometrically centered on 500 Hz. The component-frequency ratios were 1.025, 1.05, 1.1, 1.2, and 1.4, which yielded frequencies of 487–500–513 Hz, 476–500–525 Hz, 455–500–550 Hz, 417–500–600 Hz, and 357–500–700 Hz, respectively. The signal had a duration of 400 ms (50-ms cosine² rise/fall time). At the beginning of a threshold run and between trials, the listener had the option of being reminded of the three frequencies that were being evaluated currently. When two of the response box buttons were pressed simultaneously, the three current frequencies were presented (in the masking noise) sequentially at a level of 80 dB SPL. The reminding frequency sequence continued as long as the listener held the two buttons of the response box down.

In the experiment examining MLD, the masker was a low-pass noise with a frequency cutoff of 1000 Hz. In the experiment examining masking release in modulated noise, a broadband noise was square-wave modulated at a rate of either 25 Hz (modulated) or 1000 Hz (nominally unmodulated) and then digitally filtered (Trinder, 1983) to be low-pass filtered at 1000 Hz. The 1000-Hz modulation was used rather than unmodulated noise to minimize signal-path differences between the masking noise conditions. The sampling rate for the digital filter was 5.0 kHz. All noise stimuli (for both MLD and modulated noise experiments) were presented continuously at a spectrum level of 50 dB SPL. As mentioned above, we were unable to separate within-channel from across-channel masking release effects in the pitch-ranking conditions in modulated noise. We did, however collect data to obtain an approximation of the within-channel masking release for signal detection. To do this, we digitally filtered (again using a sampling frequency of 5 kHz) the unmodulated and modulated noises to bandwidths that approximated the equivalent rectangular bandwidths of the auditory filters (Moore and Glasberg, 1987) at the extreme fre-
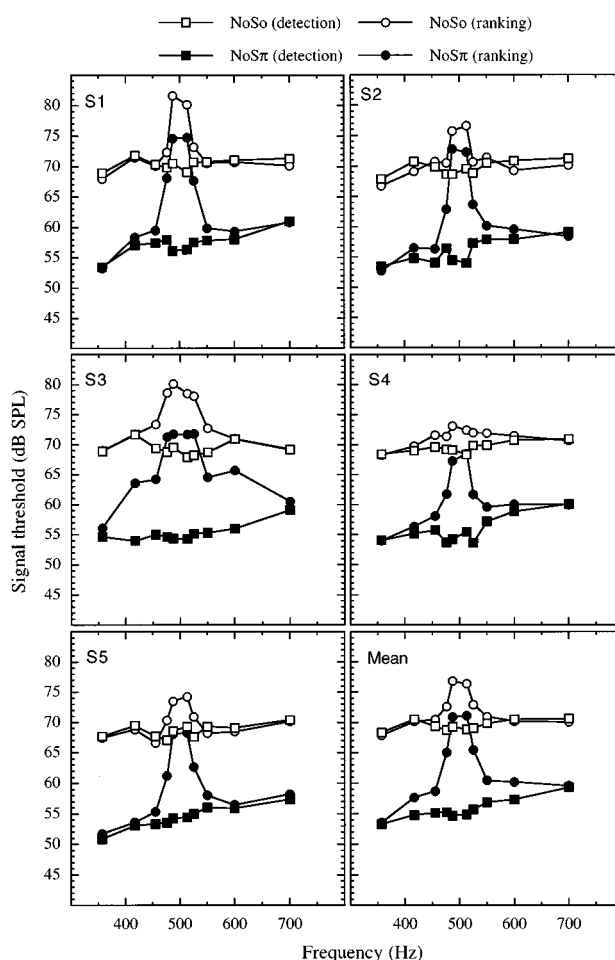
FIG. 1. Thresholds for NoSo (open symbols) and NoSπ (filled symbols) conditions for detection (squares) and pitch ranking (circles). Thresholds were determined in conditions where one of three possible signal frequencies (below, at, or above 500 Hz) was presented on a given trial. Data shown are for the signal frequencies below and above 500 Hz.

quencies investigated (357 Hz and 700 Hz). The noise bandwidth tested at 357 Hz was 64 Hz, and the noise bandwidth tested at 700 Hz was 98 Hz.

Due to listener attrition, data were incomplete for the modulated-noise conditions. Specifically listener 3 failed to provide pitch-ranking data for the 487–500–513 Hz condition, and listener 4 failed to provide detection and pitch-ranking data for the 417–500–600 Hz conditions.

II. RESULTS AND DISCUSSION

A. MLD

When considering the present results, it should be kept in mind that the data were collected in conditions where one of three possible signal frequencies could be presented on a given trial. Data for the signal frequencies presented on either side of 500 Hz are shown in Fig. 1, and data for the 500-Hz signal frequency are shown in Fig. 2. Average intra-subject standard deviations were 1.14 dB for NoSo detection conditions, 1.02 dB for NoSo ranking conditions, 1.07 dB for NoSπ detection conditions, and 1.08 dB for NoSπ ranking conditions. The mean detection results shown in Fig. 1 indicated that NoSo thresholds were relatively stable across fre-

FIG. 2. Same as Fig. 1, but for the 500-Hz component.

frequency on MLD (Hirsh, 1948). Two features can be noted concerning the average pitch-ranking thresholds shown in Figs. 1 and 2. The first is that for the highest ratio of frequency separation tested (1.4), detection and ranking thresholds were very similar to each other in magnitude, both for NoSo and NoSπ presentation. The second notable feature is that as the frequency separation between stimuli was reduced, pitch-ranking thresholds were often higher than signal detection thresholds. This was particularly the case for the NoSπ conditions, and the increase of the pitch-ranking threshold was especially steep for frequency ratios between 1.1 and 1.025. The fact that NoSπ pitch-ranking thresholds were raised more by stimulus frequency proximity than were NoSo pitch-ranking thresholds is reflected by the fact that MLDs (Tables I and II) were smaller for pitch ranking than signal detection, particularly for the sets of conditions with the smallest frequency ratios (between 1.1 and 1.025). These data indicate that masking release for pitch ranking is as large as for tone detection when the signal frequencies are widely spaced, but that such masking release for pitch ranking diminishes as the frequency spacing between the signal frequencies is reduced. These results are generally consistent with previous findings (Gebhardt *et al.*, 1971; Henning, 1973, 1991b) that have indicated relatively small MLDs for frequency discrimination, and are consistent with the hypothesis that the frequency cues are somehow more robust in NoSo than in NoSπ conditions at relatively low sensation levels.

B. Masking release in modulated noise

Figures 3 and 4 summarize individual and average threshold data for the detection and pitch-ranking conditions involving unmodulated and modulated noise. Figure 3 shows data for tones presented on either side of 500 Hz, and Fig. 4 shows data for the 500-Hz signal when presented in the context of surrounding signals at the different frequency separations. Average intrasubject standard deviations were 1.16 dB for conditions involving detection in unmodulated noise,

quency (within a range of about 2 dB), but that NoSπ thresholds gradually increased from low to high signal frequency. The average detection MLD, therefore, decreased from a value of approximately 15 dB at 357 Hz to a value of approximately 11 dB at 700 Hz (see Table I). This result is consistent with previous findings on the effect of stimulus

TABLE I. MLDs (dB) for signals presented in the context of two other signals. Data are shown for detection (d) and pitch-ranking (r) conditions.

	Signal frequencies (Hz)									
	357	417	455	476	487	513	525	550	600	700
L1(d)	15.5	14.7	12.9	11.9	14.3	12.8	13.2	13.0	13.0	10.3
L1(r)	14.7	13.1	10.6	4.2	7.0	5.4	5.5	10.7	11.4	9.3
L2(d)	14.4	15.9	15.7	12.2	14.2	15.5	11.6	12.6	12.9	12.1
L2(r)	14.0	12.6	14.4	7.6	3.0	4.3	7.0	11.2	9.7	11.7
L3(d)	14.2	17.7	14.4	14.1	15.2	13.6	13.0	13.4	14.9	10.0
L3(r)	12.9	8.1	9.1	7.4	8.4	6.8	6.3	8.2	5.3	8.7
L4(d)	14.3	13.8	13.9	15.5	14.9	12.9	16.2	12.7	11.8	10.8
L4(r)	14.2	13.4	13.5	9.6	5.9	3.9	10.3	12.3	11.4	10.7
L5(d)	16.9	16.4	14.4	13.6	14.3	14.9	12.7	13.3	13.2	13.1
L5(r)	15.8	15.2	11.2	9.1	5.4	6.0	8.2	10.2	12.1	12.0
Mean(d)	15.1	15.7	14.2	13.5	14.6	13.9	13.3	13.0	13.2	11.3
Mean(r)	14.3	12.5	11.8	7.6	5.9	5.3	7.5	10.5	10.0	10.5

TABLE II. MLDs (dB) for 500-Hz signals presented in the context of two other signals at the frequency ratios listed. Data are shown for detection (d) and pitch-ranking (r) conditions.

	Frequency ratio				
	1.025	1.05	1.1	1.2	1.4
L1(d)	12.6	12.8	14.5	13.3	12.9
L1(r)	5.3	5.3	12.2	11.7	12.1
L2(d)	15.2	12.4	14.8	14.4	15.1
L3(r)	4.0	7.8	13.4	11.1	14.7
L3(d)	13.9	14.5	13.8	15.3	14.6
L3(r)	7.1	7.7	8.6	5.7	13.2
L4(d)	14.6	15.2	12.0	12.8	14.2
L4(r)	5.6	9.3	11.7	12.3	14.0
L5(d)	13.5	14.7	14.7	15.1	13.5
L5(r)	4.5	10.2	11.6	13.9	12.4
Mean(d)	14.0	13.9	14.0	14.2	14.0
Mean(r)	5.3	8.1	11.5	11.0	13.3
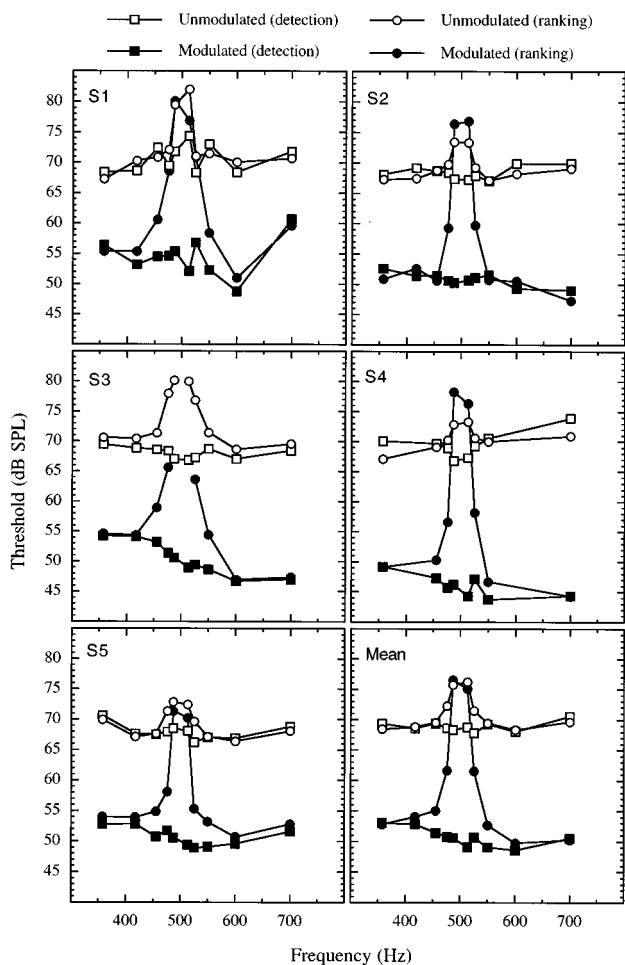
FIG. 3. Thresholds for unmodulated- (open symbols) and modulated- (filled symbols) noise conditions for detection (squares) and pitch ranking (circles). Thresholds were determined in conditions where one of three possible signal frequencies (below, at, or above 500 Hz) was presented on a given trial. Data shown are for the signal frequencies below and above 500 Hz.
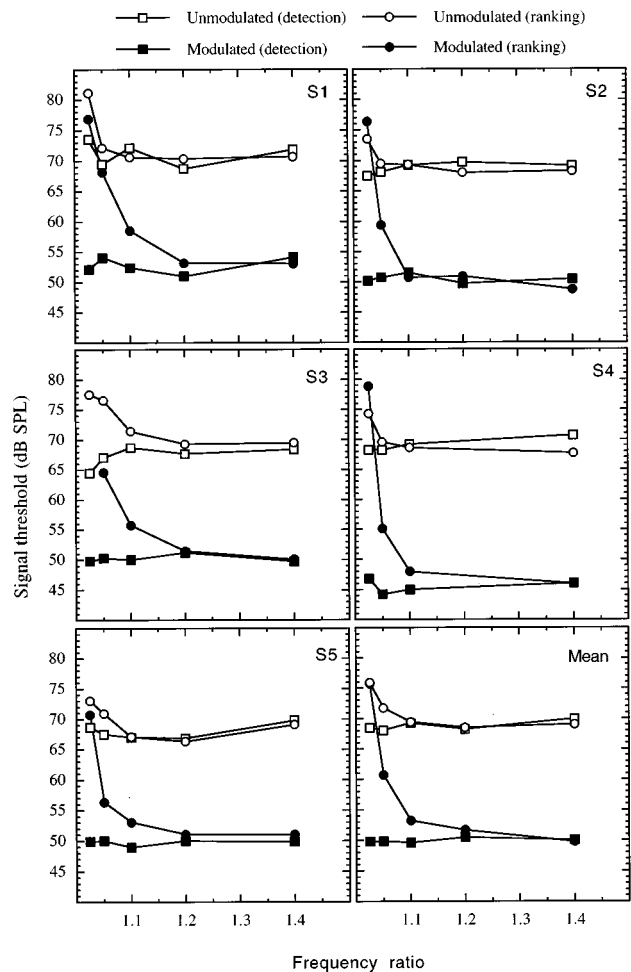
FIG. 4. Same as Fig. 3, but for the 500-Hz component.

1.12 dB for conditions involving ranking in unmodulated noise, 1.15 dB for conditions involving detection in modulated noise, and 1.30 dB for conditions involving ranking in modulated noise. The mean detection results indicated that unmodulated-noise thresholds were relatively stable across frequency. The mean detection thresholds in modulated noise were also relatively stable across frequency (within 2–3 dB), with a slight tendency for detection thresholds to decrease from low to high frequency (see Fig. 3). Trends occurring for the pitch-ranking thresholds in modulated noise were similar to those noted in the MLD conditions: for the largest frequency ratios, thresholds for detection and pitch ranking were very similar in both the unmodulated and modulated noises; for relatively small frequency ratios, pitch-ranking thresholds were higher than detection thresholds. Again, this was particularly true for the masking release condition (in this case, modulated noise), and the increase of pitch-ranking threshold was particularly steep for frequency ratios between 1.1 and 1.1025. This pattern of data resulted in the unmodulated-modulated difference (Tables III and IV) being smaller for pitch ranking than detection, particularly for the sets of conditions with the smallest frequency ratios (between 1.1 and 1.025). Indeed, the average masking release was essentially zero at the 1.025 ratio. For listeners 2 and 4, the pitch-ranking threshold at the 1.025 ratio was actually higher in modulated than in unmodulated noise. As

TABLE III. Unmodulated-noise threshold minus modulated-noise threshold (dB) for signals presented in the context of two other signals. Data are shown for detection (d) and pitch-ranking (r) conditions.

	Signal frequencies (Hz)									
	357	417	455	476	487	513	525	550	600	700
L1(d)	12.0	15.4	17.9	14.8	16.4	22.2	11.5	20.7	19.6	11.2
L1(r)	11.9	14.9	10.2	3.5	−0.7	5.1	0.2	13.1	19.1	11.0
L2(d)	15.5	17.8	17.2	17.8	17.1	16.6	16.8	15.5	20.6	20.9
L2(r)	16.4	14.8	18.1	10.5	−3.0	−3.5	9.5	16.4	17.6	21.8
L3(d)	15.2	14.8	15.4	17.0	16.4	17.9	17.8	20.0	20.4	21.4
L3(r)	16.0	16.1	12.4	12.3	⋯	⋯	13.2	17.0	21.7	22.2
L4(d)	20.9	⋯	22.3	23.2	20.6	23.0	22.1	26.8	⋯	29.5
L4(r)	18.0	⋯	18.8	13.7	−5.0	−3.0	12.4	23.3	⋯	26.6
L5(d)	17.9	14.8	16.7	16.2	17.9	18.6	17.3	17.9	17.3	17.2
L5(r)	16.0	13.2	12.7	13.2	1.5	2.2	14.3	13.9	15.7	15.3
Mean(d)	16.3	15.7	17.9	17.8	17.7	19.7	17.1	20.2	19.5	20.0
Mean(r)	15.7	14.8	14.4	10.6	−1.8	0.2	9.9	16.7	18.5	19.4

for MLD, the modulated-noise data indicate that masking release for pitch ranking was as large as for signal detection when the target tones were very widely spaced, but diminished as the frequency spacing between the tonal set was reduced. Thus a new finding of the present study is that suprathreshold masking release for pitch ranking is relatively small in modulated noise when frequency spacing between components is relatively small. This would suggest that frequency discrimination would also be relatively poor in modulated noise. One dissimilarity between the modulated-noise results and the MLD results is that whereas a small MLD for pitch ranking remained at the narrowest frequency spacing tested (1.025), there was essentially no masking release in modulated noise at this frequency separation.

The threshold differences between unmodulated and modulated noise were relatively small for the 64-Hz-wide noisebands centered on 357 Hz and the 98-Hz-wide noise bands centered on 700 Hz. The average thresholds (over five listeners) for the 64-Hz-wide noise bands were 66.5 dB SPL for unmodulated noise and 62.2 dB SPL for modulated noise (difference of 4.3 dB), and the average thresholds for the 98-Hz-wide noise bands were 67.3 dB SPL for unmodulated noise and 60.1 dB SPL for modulated noise (difference of 7.2 dB). These results would imply that part of the masking release obtained in the modulated noise of the main experiment was due to across-channel masking release. As noted in the Introduction, the masking release in modulated noise obtained in this study is probably due both to within-channel and to across-channel processes. In this regard, it is interesting that, at least at the smallest frequency ratio tested, essentially no masking release occurred for pitch ranking. This would suggest that neither the within-channel nor the across-channel cue was sufficiently robust to provide a masking release for pitch ranking at the 1.025 frequency ratio.

TABLE IV. Unmodulated-noise threshold minus modulated-noise threshold (dB) for 500-Hz signals presented in the context of two other signals at the frequency ratios listed. Data are shown for detection (d) and pitch-ranking (r) conditions.

	Frequency ratio				
	1.025	1.05	1.1	1.2	1.4
L1(d)	21.4	15.4	19.7	17.7	17.7
L1(r)	4.3	4.0	12.1	17.2	17.6
L2(d)	17.2	17.4	17.7	20.0	18.6
L2(r)	−2.8	10.1	18.6	17.1	19.5
L3(d)	14.6	16.7	18.6	16.5	18.6
L3(r)	⋯	12.0	15.6	17.8	19.4
L4(d)	21.4	24.0	24.2	⋯	24.6
L4(r)	−4.6	14.5	20.6	⋯	21.7
L5(d)	18.8	17.5	18.0	16.9	19.9
L5(r)	2.3	14.5	14.0	15.2	18.0
Mean(d)	18.7	18.2	19.6	17.8	19.9
Mean(r)	−0.2	11.0	16.2	16.8	19.2

III. GENERAL DISCUSSION

A. Relation to past MLD findings

The present pitch-ranking conditions can be considered to be highly similar to tasks of frequency discrimination. The result indicating that MLD for pitch ranking decreases as the frequency separation between the signals decreases is therefore in agreement with past studies that have shown that frequency discrimination is often poorer in NoSπ than NoSo conditions when performance is compared at relatively low sensation levels. Henning (1990,1991b) noted that such a result is not in good agreement with the equalization–cancellation (EC) model of MLD (Durlach, 1963), because operation of the EC mechanism should result in a simple improvement in the signal-to-noise ratio for NoSπ processing. With such an improvement, discrimination performance should be similar in NoSo and NoSπ presentations at like sensation levels.

B. Performance in modulated-noise and supplementary experiment

The present finding of relatively poor pitch ranking for tones presented in modulated noise is a new result. However, the mechanisms responsible for this result are probably closely related to those underlying poor gap detection (Hall and Grose, 1992) and poor amplitude discrimination (Hall and Grose, 1995) in comodulated narrow bands of noise. Hall and Grose (1995) suggested that the dip-listening model of CMR (Buus, 1985) has the same problem accounting for poor amplitude discrimination in comodulated noise as the EC model has in accounting for poor amplitude discrimination performance in NoSπ conditions. In the dip-listening model, comodulated noise remote from the signal frequency is used by the auditory system to identify low energy or dip regions in the waveform of the on-signal masker. When the envelope of the comodulated noise is relatively high, low weight is given, and when the envelope of the flanking band is low, high weight is given. This rule essentially improves the signal-to-noise ratio. If the primary result of the dip-listening mechanism is simply to improve the signal-to-noise ratio, then the poor performance on signal discriminations or pitch ranking at low signal-to-noise ratios needs to be addressed. Below we consider an interpretation of a dip-listening strategy that may be compatible with the finding of relatively poor discrimination/pitch-ranking of signals in comodulated noise.

One factor that might be considered in relating the dip-listening model to discrimination/pitch-ranking performance in modulated noise concerns the fact that the signal is assumed to be extracted from intermittent epochs where the noise envelope is relatively low, separated by epochs where the noise envelope is relatively high. For example, in the present experiment, the noise was modulated such that dip epochs of 20 ms were separated by noise epochs of 20-ms duration. One possible interpretation of the data for pitch ranking in modulated noise is that poor performance occurs because the task required synthesis of frequency information from several short signal bursts that were separated in time. Although it is known that reductions in signal duration result in decreases in frequency discrimination performance (Moore, 1973; Hall and Wood, 1984; Freyman and Nelson, 1986), it is not as clear what might be expected when several short bursts, separated in time, are available for analysis. It would be informative to determine the difference in performance between conditions where the signal was a single, relatively long duration tone, and conditions where the signal was a series of relatively short bursts that are separated in time. We therefore conducted a simple supplementary experiment where we actually presented multiple, short-duration stimuli in noise. A single, new listener was first run on the same modulated-noise conditions of the main experiment. Then, conditions were run in which signals were presented as a train of 10 tone bursts, having an overall duration of 400 ms. The temporal shaping of the stimulus was based on considerations related to the main experimental conditions where a long-duration signal was presented in the modulated masking noise. If performance is based on signal information in the masker dips, then it may be assumed that
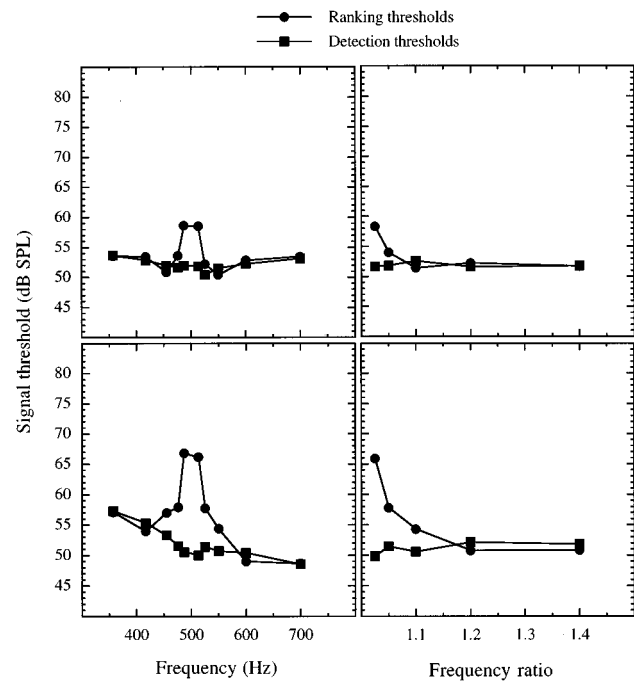
FIG. 5. Thresholds for detection (squares) and pitch ranking (circles). Thresholds were determined in conditions where one of three possible signal frequencies (below, at, or above 500 Hz) was presented on a given trial. Data in the left column are for the signal frequencies below and above 500 Hz, and data in the right column are for the 500-Hz signals at the different frequency ratios tested. The data in the upper row are for a train of short signals presented in an unmodulated noise. The data in the lower row are for a 400-ms tone presented in modulated noise.

the forward masking (Elliott, 1962) of the signal decreases over the duration of the masker dip. We therefore manipulated stimulus temporal shaping to take into account the fact that signal information was not optimally available over the entire 20 ms of a masker dip: The stimulus duration was reduced to 16 ms, and a relatively protracted rise time was used. The cosine² rise time for each burst was 10 ms and the fall time was 5 ms, with a steady-state portion of approximately 1 ms. The time interval between the offset of a tone burst and the onset of the following tone burst was approximately 24 ms. In the multiple tone burst condition, the masking noise level was adjusted so that the detection threshold would be about the same as occurred for detection of the 400-ms tone in the 25-Hz modulated noise. A masking noise pressure spectrum level of 26 dB was used for this purpose.

The results of the supplementary experiment are shown in Fig. 5. The data obtained in modulated noise (lower panels) agreed broadly with those of the listeners in the main experiment: Detection and pitch-ranking thresholds were similar for the wider frequency separation, but pitch-ranking thresholds were considerably higher than detection thresholds for the relatively narrow frequency separations. The pattern of results for the tone bursts (upper panels) was somewhat different in detail, however. Although pitch-ranking thresholds increased as the frequency separation between tones decreased, the increases were smaller than those occurring for long-duration tones presented in modulated noise. This result would suggest that it is unlikely that the poor pitch-ranking performance found in the main experiment in

modulated noise was due entirely to the fact that frequency information had to be analyzed across temporally separated, relatively short, temporal epochs. With this interpretation, the relatively poor pitch-ranking performance in modulated noise would remain as a problem for the dip-listening model of CMR. There is at least one caveat to this interpretation, however. Although we attempted to choose a signal duration that would mimic the duration available in the modulated-noise conditions of the main experiment, the actual effective duration in the main experiment may have been longer or shorter. It is possible that the results of the supplementary experiment did not match those of the main experiment because the effective signal duration in the main experiment was shorter than the signal duration used in the supplementary experiment.

C. Relation between results for MLD and masking release in modulated noise

There were both similarities and differences between the patterns of results obtained in the binaural and monaural masking release conditions. The most striking similarity was that performance in both types of masking release conditions decreased steeply when the frequency ratio was between 1.1 and 1.025. Although the steepness of performance decline was less remarkable in the nonmasking release conditions, it is noteworthy that pitch-ranking thresholds here were also most markedly elevated between the 1.1 and 1.025 ratios. The significance of this pattern may be that the same cue or cues are used to estimate stimulus pitch in all of the conditions tested.

The pitch-ranking masking-release results for MLD and for modulated-noise conditions did differ in terms of the magnitude of the reduction in masking release that occurred as the frequency separation decreased. The most prominent difference was that, over the range of frequency ratios tested, the average MLD ranged from a maximum of about 14 dB for the 1.4 ratio to a minimum of about 5 dB for the 1.025 ratio, whereas the average masking release for modulated noise ranged from a maximum of about 19 dB for the 1.4 ratio to a minimum of about 0 dB for the 1.025 ratio. Thus whereas a pitch-ranking masking release for MLD remained at the smallest ratio tested, there was no pitch-ranking masking release for modulated noise at this frequency ratio. The greatest difference between the MLD/modulated-noise functions occurred between the ratios of 1.1 and 1.025. Here the average pitch-ranking MLD reduced by about 6 dB, while the masking release in modulated noise reduced by about 16 dB. This would imply that the information important for pitch ranking is more coarsely represented in modulated noise than in the $NoS\pi$ conditions at low signal-to-noise ratios, a factor that becomes relevant for performance only when the task is relatively difficult (when the frequency ratio is relatively small).

IV. CONCLUSIONS

(1) Masking release for pitch ranking was as large as for signal detection when the frequency ratio between adjacent components was 1.2 or higher, both for masking release re-

lated to binaural analysis and masking release related to noise modulation. However, masking release for pitch ranking was considerably smaller than masking release for signal detection when the frequency ratio was 1.1 or smaller.

(2) It is unlikely that poor pitch-ranking performance in modulated noise is due solely to a difficulty in determining pitch on the basis of relatively short-duration representations of the signal that occur in the masker dips.

(3) Information important for the determination of pitch appears to be more coarsely represented in modulated noise than under $NoS\pi$ conditions, at similar levels above masked detection threshold.

[1]Previous research has indicated that signal frequency uncertainty effects in CMR conditions (Grose and Hall, 1990) are only 1 or 2 dB, similar to the size of the uncertainty effect found in traditional masking conditions (Green, 1961). Furthermore, pilot data on both MLD and modulated-noise conditions indicated effects of signal frequency uncertainty that were similar in magnitude for the masking release and nonmasking release conditions. Therefore, the masking release reported here for detection, where signal frequency was uncertain, should be of similar magnitude to the values of masking release typically found under fixed-frequency stimulation.

[2]In pilot data on pitch ranking, we also examined procedures where the attenuation of each component could vary independently, with each component tracked by a separate staircase. In this procedure, some listeners appeared to use a strategy based upon signal audibility. By the random nature of the adaptive task, the three different frequencies might become identifiable on the basis of different degrees of audibility. If, for example, the middle component were driven to a very low level where it was rarely audible, a listener might continue to rank its pitch correctly not on the basis of its pitch, but on the basis that no stimulus was heard on that trial.

Buus, S. (**1985**). ''Release from masking caused by envelope fluctuations,'' J. Acoust. Soc. Am. **78**, 1958–1965.

Carlyon, R. P., Buus, S., and Florentine, M. (**1989**). ''Comodulation masking release for three types of modulators as a function of modulation rate,'' Hearing Res. **42**, 37–46.

Durlach, N. I. (**1963**). ''Equalization and cancellation theory of binaural masking-level differences,'' J. Acoust. Soc. Am. **35**, 1206–1218.

Elliott, L. L. (**1962**). ''Backward and forward masking of probe tones of different frequencies,'' J. Acoust. Soc. Am. **34**, 1116–1117.

Freyman, R. L., and Nelson, D. A. (**1986**). ''Frequency discrimination as a function of tonal duration and excitation-pattern slopes in normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **79**, 1034–1044.

Gebhardt, C. J., Goldstein, D. P., and Robertson, R. M. (**1971**). ''Frequency discrimination and the MLD,'' J. Acoust. Soc. Am. **51**, 1228–1232.

Green, D. M. (**1961**). ''Detection of auditory sinusoids of uncertain frequency,'' J. Acoust. Soc. Am. **33**, 897–903.

Grose, J. H., and Hall, J. W. (**1990**). ''The effect of signal-frequency uncertainty on comodulation masking-release,'' J. Acoust. Soc. Am. **87**, 1272–1277.

Grose, J. H., and Hall, J. W. (**1992**). ''Comodulation masking release for speech stimuli,'' J. Acoust. Soc. Am. **91**, 1042–1050.

Hall, J. W., and Grose, J. H. (**1992**). ''Masking release for gap detection,'' Philos. Trans. R. Soc. London, Ser. B **336**, 331–337.

Hall, J. W., and Grose, J. H. (**1995**). ''Amplitude discrimination in masking release paradigms,'' J. Acoust. Soc. Am. **98**, 847–852.

Hall, J. W., and Wood, E. J. (**1984**). ''Stimulus duration and frequency discrimination for normal-hearing and hearing-impaired subjects,'' J. Speech Hear. Res. **27**, 252–256.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). ''Detection in noise by spectro-temporal pattern analysis,'' J. Acoust. Soc. Am. **76**, 50–56.

Henning, G. B. (**1973**). ''Effect of interaural phase on frequency and amplitude discrimination,'' J. Acoust. Soc. Am. **54**, 1160–1178.

Henning, G. B. (**1990**). ''The effect of interaural phase on frequency discrimination with broad- and narrow-band maskers,'' Hearing Res. **48**, 195–200.

Henning, G. B. (**1991a**). ''Binaural amplitude discrimination and the binaural masking-level difference,'' Hearing Res. **55**, 188–194.

Henning, G. B. (**1991b**). ''Frequency discrimination, amplitude discrimination, and the binaural masking-level difference: Some anomalous results,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford).

Hirsh, I. J. (**1948**). ''Influence of interaural phase on interaural summation and inhibition,'' J. Acoust. Soc. Am. **20**, 536–544.

Humes, L. E. (**1990**). ''Masking of tone bursts by modulated noise in normal, noise-masked normal and hearing-impaired listeners,'' J. Speech Hear. Res. **33**, 3–8.

Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (**1952**). ''The masking of tones by white noise as a function of the interaural phases of both components: I. 500 cycles,'' J. Acoust. Soc. Am. **24**, 523–527.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Levitt, H., and Rabiner, L. R. (**1967**). ''Binaural release from masking for speech and gain intelligibility,'' J. Acoust. Soc. Am. **42**, 601–608.

Moore, B. C. J. (**1973**). ''Frequency difference limens for short-duration tones,'' J. Acoust. Soc. Am. **54**, 610–619.

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,'' Hearing Res. **28**, 209–225.

Trinder, J. R. (**1983**). ''Hardware-software configuration for high performance digital filtering in real time,'' Proc. IEEE Int. Conf. Acoust., Speech Signal Process. **2**, 687–690.

Zwicker, E., and Schorn, K. (**1982**). ''Temporal resolution in hard of hearing patients,'' Audiology **21**, 474–492.

Frequency discrimination of stylized synthetic vowels with a single formant

Johannes Lyzenga[a] and J. Wiebe Horst
Department of Otorhinolaryngology/Audiology, University Hospital Groningen, P.O. Box 30.001,
9700RB Groningen, The Netherlands

Just-noticeable differences (jnd's) in the center frequency of bandlimited harmonic complexes were measured for normal-hearing subjects. A triangular and a rounded spectral envelope were used. The center frequency ranged from 500 to 600 Hz in a region representing the first formant of vowels, and from 2000 to 2100 Hz in a second formant region. The slope of the spectral envelope was either 50 or 100 dB/oct for the first formant region and 100 or 200 dB/oct for the second formant region. For the fundamental frequency of the complexes 100 and 200 Hz were used. The jnd's were determined for various phase relations between the individual components of the complexes. For comparison we also determined jnd's for a Gaussian white noise that was filtered with the same spectral envelopes as the harmonic complexes. A three-interval, three-alternative forced-choice task was used. All measurements were performed with roving stimulus level. The jnd's found for center frequencies that were halfway between two harmonics were smaller than those found for center frequencies that coincided with a harmonic. The jnd's for the noise bands were mostly between those of the two aforementioned groups. Except for a small group of stimuli, the phase relations had little effect on the jnd's. The majority of the results for both the harmonic and the noise band stimuli can be described by a model using a spectral profile comparison. Most of the remaining data can be explained in the temporal domain from changes in the temporal envelope of the stimuli. © *1997 Acoustical Society of America.* [S0001-4966(97)00609-7]

PACS numbers: 43.66.Fe, 43.66.Jh, 43.71.Es [WJ]

INTRODUCTION

In order to investigate the ability of normal-hearing listeners to discriminate spectral changes in vowels, we have measured frequency discrimination for bandlimited tones with a spectrum resembling that of a single vowel-formant (resonance peak). In vowels of natural speech a number of formants can be recognized, the first two or three of which characterize the vowel. The ability to hear changes in the character of a vowel during pronunciation depends on the sensitivity of the ear to changes in these formant frequencies. To gain understanding of the sensitivity for changes in formants, it is a logical first step to study the sensitivity for center frequency changes in tones with just a single peak in the spectrum. After investigating frequency discrimination for such tones with different center frequencies, two or more of these tones can be used together to achieve a spectrum more closely approximating that of a vowel. In a later paper we will report on the frequency discrimination of the ''formants'' of such combined tones, in relation to the known frequency discrimination of the constituting tones.

In the present paper we present an investigation of frequency discrimination for stimuli in a first and in a second formant region. We used two shapes for the spectral envelope shape: a triangular shape in a first experiment, and a more natural shape in a second experiment. The triangular envelope shape was used before by Lyzenga and Horst (1995) for stimuli in the second formant region with a 100-Hz fundamental and a sine phase relation. We used the

fundamentals of 100 and 200 Hz, representative for male and female speakers, respectively. Lyzenga and Horst used three values for the spectral slopes; from these three values we chose the two most realistic values for the present stimuli. Lyzenga and Horst found that the position of the center frequencies relative to the harmonics was an important parameter, so we chose two dissimilar values for this parameter.

The role of the phase spectrum in hearing has been studied in several investigations. In general, it appears that differences in phase spectrum can be heard when they lead to audible differences in temporal envelopes (Duifhuis, 1972). Effects have been reported, e.g., in masking (Duifhus, 1970), pitch (Ritsma and Engel, 1964; Moore, 1977), pitch perceptibility (Bilsen, 1973; Lundeen and Small, 1984), timbre (Plomp and Steeneken, 1969), and frequency discrimination (Hoekstra and Ritsma, 1977; Moore and Peters, 1992). The effects are often not strong, and occasionally no effects were found (Patterson, 1973). In recordings of responses in single auditory-nerve fibers, a strong influence of relative phase on the period histograms (Horst *et al.*, 1990) and interspike-interval histograms can be found (Horst *et al.*, 1992). Goldstein and Srulovicz (1977) have suggested that interspike intervals play an important role in frequency discrimination. By implication, variation of stimulus phase spectra might influence frequency discrimination. Therefore we used various phase spectra in the present investigation to use the occurrence of phase effects as an indicator of the involvement of temporal discrimination mechanisms.

Lyzenga and Horst (1995) measured jnd's for bandlimited harmonic tones with a triangular and a trapezoidal spectral envelope (on a log-log scale). They used center frequencies near 2 kHz and a fundamental of 100 Hz; therefore the

[a] Current affiliation: MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom.

components of their stimuli were not resolved by the auditory periphery. The stimulus components were added in sine phase. They used three values for the spectral slope (100, 200, and 400 dB/oct). Near 2 kHz they chose five to nine center frequencies in such a way that the center frequency was either located at a harmonic, or at a distance of one-quarter, one-half, or three-quarters of the fundamental frequency above a harmonic. The distance between the center frequency and the harmonics was a major influence on the jnd's, and divided the jnd's into two groups: The smallest jnd's under each condition were found when the center frequency was halfway between two harmonics for the triangular envelope and when it coincided with a harmonic for the trapezoidal envelope, i.e., when the largest changing components were of approximately equal level. The jnd's under these conditions showed an inverse proportionality with the slope. A similar inverse proportionality was found by Horst *et al.* (1984). They used a proportional change in the center and the fundamental frequency, whereas these two parameters were independent in the experiments of Lyzenga and Horst. Because of the correspondence between both the spectral envelopes and the jnd's of both studies, Lyzenga and Horst suggested a modified version of the place model (see below for a detailed description) as a possible mechanism underlying the discrimination for these jnd's. In this modified place model, the excitation differences caused by both the low- and the high-frequency slopes of the stimuli were combined to increase sensitivity (see Fig. 10 of Lyzenga and Horst, 1995). This combination was proposed to be stimulus dependent, so as to work well for the triangular stimuli, but poorly for the trapezoidal stimuli, where the upper and lower flanks were separated by a central plateau. Their data, however, were not sufficiently extensive to decide whether the discrimination mechanism was strictly spectral, or whether a temporal mechanism was involved, based on waveform changes within one frequency channel of the auditory system. The jnd's for the remaining conditions in the study of Lyzenga and Horst were not inversely proportional to the slope. For the triangular envelope, however, these jnd's correlated very well with the amplitude modulation depth: The modulation power difference between just-noticeably different stimuli showed a linear correlation with the initial modulation depth. For the trapezoidal envelope the jnd's correlated reasonably well with the modified place model. For these stimuli this model is in fact equivalent to the original (i.e., nonmodified) place model.

The modified place model used by Lyzenga and Horst (1995) consists of a linear filter bank followed by a detector of level differences. The filter bank consists of 3400 roex filters (rounded exponential; Moore and Glasberg, 1987) with a Q of 5:

$$|H_{cf}(f)| = \sqrt{[(1+pg)\exp(-pg)]}, \qquad (1)$$

where $g = |f - cf|/cf$, and $p = 6.849Q$. The spacing of the center frequencies of these filters is logarithmic above, and linear below, 800 Hz. This linear relation seems to be in contrast with the constant Q of the filters. These relations, however, were taken for practical reasons. For this low-frequency range they were not of critical influence on the

predicted jnd's, as the stimulus components were to a large extent resolved for the fundamental frequencies we used. To form an excitation pattern, the outputs of the filters of all channels are converted to a level L_n in dB (Fletcher, 1940). As a representation of the absolute threshold, a noise floor with a power of 1 (i.e., a level of 0 dB) is added to the output power of each filter before the conversion to dB:

$$L_n = 10 \cdot {}^{10}\log\left[1 + \int_{-\infty}^{\infty} H_{cf}(f)^2 \cdot S(f)^2 \, df\right], \qquad (2)$$

where $S(f)$ denotes the amplitude spectrum of the stimulus. The level difference detector compares the excitation patterns for two tones (Zwicker, 1970; Patterson and Moore, 1986). This detector has been modified for the application of the model to experiments with roving stimulus levels. In the original place models this detector judges two tones to be perceptually different when somewhere a difference of 1 dB can be found between the two excitations, regardless of the sign of this difference. One consequence of this strategy is that an increase of 1 dB in the level of a stimulus leads to a detectable difference, which is in accordance with the often reported 1-dB overall level jnd. In experiments with roving stimulus levels, subjects have to ignore overall level information, so, a different strategy is needed. The strategy chosen by Lyzenga and Horst (1995) is to roughly equalize the excitation levels (by matching their overall levels), then to sum the maximum positive excitation difference and the maximum value of the absolute negative difference (as indicated by the arrow in the upper panel of Fig. 10, Lyzenga and Horst, 1995), and lastly, to compare this sum with the detection threshold.

Alongside their modified place model, Lyzenga and Horst (1995) used the amplitude modulation depth of the stimuli to explain the jnd's for some of their stimuli. To verify their findings, an extended group of such stimuli is used in the present study. Alternative explanations may be found in differences in the frequency modulation depth of just-noticeably different stimuli, or in differences in their EWAIF or IWAIF values (the envelope-weighted or intensity-weighted averaged instantaneous frequency; Feth, 1974). These possibilities will be considered in the General Discussion.

In the present paper, we use both the triangular spectral envelope and a broader, more natural envelope shape. For this natural shape we chose the specifications of the Klatt synthesizer (Klatt, 1980). With one formant peak, this shape spreads out over a broad spectral band at low component levels, with a sharp peak in the component levels near the formant frequency. Stimuli with comparable spectral shapes have been used by Stevens (1952) and Gagné and Zurek (1988). Stevens studied just-noticeable differences (jnd's) for changes in the frequency of a single exponentially damped wave. The spectrum of such a tone resembles the spectral envelope of one formant of a vowel. Unfortunately, the effective duration of these tones depends on the bandwidth, so, some care has to be taken when comparing these results directly with frequency jnd's of stimuli with fixed duration. Gagné and Zurek used RLC-type resonance filters to band-limit harmonic sequences that had either a 100- or 250-Hz

fundamental. They used center frequencies from 300 to 800 Hz, so, in contrast to the stimuli used by Lyzenga and Horst (1995), the components of their stimuli were well resolved. Gagné and Zurek found that their jnd's corresponded reasonably well with the expectations of a place model. In their investigation, and in that of Stevens (1952), jnd's were found to decrease with increasing slope, from roughly 4% down to 0.6%, but less steeply than was found by Horst *et al.* (1984). (The jnd's of Gagné and Zurek are corrected in value to correspond with a d' of 1, as are all jnd's in the present study.) The jnd's of the former two studies corresponded roughly with those found by Lyzenga and Horst for both the triangular and the trapezoidal spectral slope, under conditions where the largest changing components were of unequal level.

To enable a comparison of jnd's for harmonic stimuli with those for stochastic stimuli, we investigate the jnd for Gaussian white noises that are filtered with the spectral envelopes of the harmonic stimuli. These stochastic stimuli resemble single-formant ''vowels'' from whispered speech. For comparable stimuli, jnd's were investigated by Michaels (1957), Moore (1973), and Gagné and Zurek (1988). Gagné and Zurek found for filtered white-noise bands with a center frequency of 2000 Hz that jnd's (corrected in value to correspond with a d' of 1) decreased from 3% to 0.6% as the slope increased from 5 to 180 dB/oct. Michaels and Moore measured frequency difference limens for narrow bands of noise with the bandwidth of the spectral envelope as a parameter. Michaels varied the bandwidth by adjusting the Q factor of an analog filter with a center frequency of 800 Hz. Moore directly varied the bandwidth of a rectangular spectral envelope with a center frequency of 2 kHz. They both found that jnd's decreased from 0.8% to approximately 0.4% when decreasing the bandwidth from 64 to 12 Hz (Michaels), or 16 Hz (Moore). For smaller bandwidths the jnd's remained between 0.3% and 0.4%. So, the jnd's appear to depend on the bandwidth rather than on the shape of the spectral envelope. Both researchers proposed a temporal model for the description of their data. With the same subjects and apparatus both found a constant-level pure-tone jnd of 0.2%.

Lyzenga and Horst (1995) performed their experiments with both a roving and a nonroving measurement paradigm. They found a roughly constant ratio between the jnd's for these two paradigms; the jnd's for the roving condition were typically a factor of 1.5 larger. They argued that this difference was due to the increased stimulus uncertainty of the roving-level condition. Correcting the results of the roving-level condition with this factor then yields identical jnd's for both measurement paradigms. So, even though there is a constant difference between the jnd's, the discrimination mechanism can still be the same. Because of the disturbance of overall-level information under the roving-level condition, this mechanism then has to be of a kind that ignores overall level.

This study presents an investigation of formant-frequency discrimination in which we extended the stimuli used by Lyzenga and Horst (1995) with an extra formant region, fundamental frequency, spectral shape, and various phase spectra. In this way we could further investigate the

functionality of the models they proposed, and in addition make the transition from a stylized spectral shape to more natural spectral shape.

I. GENERAL METHODS

A. Stimuli

Two experiments were carried out using bandlimited harmonic complexes as stimuli. In experiment 1 the shape of the spectral envelope was triangular on a log-log scale. In experiment 2 we used a rounded envelope shape, equal to the shape of a single formant as generated by the digital Klatt synthesizer (Klatt, 1980), which we called the Klatt envelope. For both envelope shapes two values for the fundamental frequency (F_0) of the harmonic complexes were used: 100 and 200 Hz. We used two ranges of center frequency (F_C) values: one in the region of the first formant and one in the region of the second formant. Two values for the slope (G) were employed for both formant regions: 50 and 100 dB/oct for the first region and 100 and 200 dB/oct for the second region. For the Klatt envelopes, these slope values corresponded with the steepest parts of the spectral slopes. Under each condition jnd's were measured for two center frequencies: one at a harmonic component, and one halfway between two harmonic components. In the first formant region these center frequencies were 500 and either 550 or 600 Hz, and in the second formant region they were 2000 and either 2050 or 2100 Hz. In experiment 1 we used two phase relations between the harmonics of the complexes: sine phase, and a random phase. For one subject, jnd's were also investigated for the negative Schroeder-phase relation (Schroeder, 1970), which produces stimuli with a maximally flat temporal envelope. In experiment 2 we used the same three phase relations, and in addition we used the phase relation generated by the digital Klatt synthesizer (hereafter called the Klatt-phase relation). Figure 1 gives the stimulus spectra for both spectral envelope shapes. We also investigated the jnd for a Gaussian white noise that was filtered with the spectral envelopes shown in Fig. 1.

The stimulus generation and presentation procedures were the same as those described by Lyzenga and Horst (1995) for the roving level condition. For both the triangular and the Klatt envelopes, the spectral envelope shape was calculated first, after which the harmonic components were added with the appropriate amplitude and phase. For the noise stimuli and the stimuli with random-phase relations, three sets of stimuli were made with different random relations. During the measurements each stimulus was picked at random from one of these three sets. In this way the correlation between the stimuli was reduced.

B. Procedure

The jnd's were measured using an adaptive three-interval, three-alternative forced-choice method (3IFC). Before the actual jnd measurements, the absolute threshold of the reference tone was estimated, after which stimuli were presented at a level about 40 dB above this threshold. We added a pink background noise to the stimuli at a level of -40 dB relative to the stimuli, when measured with a
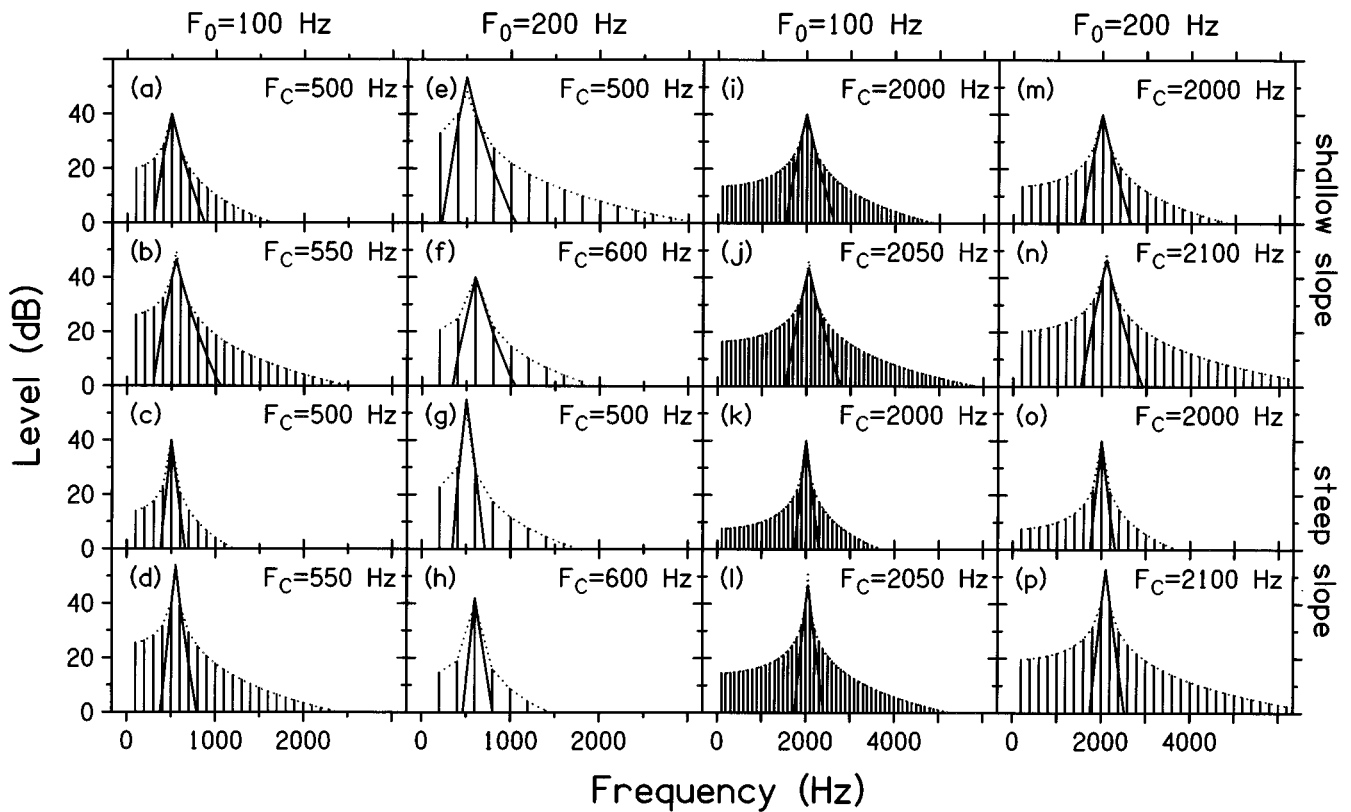
FIG. 1. The spectra of the stimuli used. The first two columns show the stimuli for the first, and the last two columns for the second formant region. The upper two rows display the stimuli with the shallow slopes: 50 dB/oct for the first, and 100 dB/oct for the second formant region, and the lower two rows those with the steep slopes: 100 dB/oct for the first, and 200 dB/oct for the second formant region. The dotted lines indicate the spectral envelopes according to the Klatt synthesizer (1980), the slanting solid lines indicate the triangular spectral envelopes.

120-Hz bandwidth for the first formant region, and with a 240-Hz bandwidth for the second formant region. These bandwidth values approximate the critical bands for both formant regions (Scharf, 1970); therefore, this procedure produced background noise levels close to the absolute thresholds. We used a "roving level" (RL) condition in which the levels of the stimuli were randomized, around one fixed level value, within trials over a 10-dB range in 0.5-dB steps (Henning, 1966). Prior to randomization, the sets of stimuli (containing a reference and twelve targets) used in each jnd estimation were of roughly constant loudness. Therefore it is reasonable to expect that any remaining overall-level information, that the subjects might have used as a discrimination cue, will have been disturbed by the 10-dB rove. The background noise was varied in level along with the roved stimulus, keeping the signal-to-noise ratio in the stimulus constant.

The adaptive procedure by which the jnd's were estimated was equal to the one used, and described in detail, by Lyzenga and Horst (1995). In short, subjects were asked to identify the odd tone in a series containing two reference tones and one target tone with a higher center frequency. They were given immediate feedback. The frequency difference between the target and the reference tones was adapted according to decision rules that were chosen so that the procedure converged at 63% correct responses, which corresponds to a d' of 1 for the 3IFC paradigm. Data were collected until the direction of the center frequency adaption was reversed five times. On the average, one jnd measure-

ment contained about 71 trials. The whole set of jnd's was measured three times (involving at least 200 trials per jnd), in a pseudorandom order for the two groups of stimuli with either the triangular or the Klatt envelope. The jnd's were estimated from the averaged scores with the same algorithm as used by Lyzenga and Horst (1995). Since the subjects' bias toward one of the three signal intervals was found to be very small, it has been neglected in the calculations.

C. Subjects

Six normal-hearing subjects participated in the experiments. All were adults, four female and two male, with ages ranging from 24 to 47 years. Four subjects participated in experiment 1; three of these subjects performed a series of jnd estimations for two phase relations and for the filtered noise bands. For subject JL the Schroeder phase was also included. Six subjects participated in experiment 2. All subjects performed a series of jnd estimates for two phase relations and for the filtered noise bands, except for subject JWH who participated for one phase relation and for the noise bands, and for subject JTB who participated for two phase relations. Of these six subjects, five had participated in frequency discrimination experiments before. The new subject was trained before data collection. For all six subjects no improvements in the scores were observed during the course of the measurements.

II. EXPERIMENT 1: THE TRIANGULAR SPECTRAL ENVELOPE

The results of the first experiment are presented in Fig. 2. Each row contains the just-noticeable differences, measured for one region of the center frequency and one value of the slope, in four panels with individual results. The columns contain the jnd's for the four combinations of formant frequency and slope. Each panel contains jnd's for two values of the center frequency: one coinciding with and one halfway between the signal components, denoted as /I\ and /II\, respectively. The smaller value of these pairs of center frequencies is plotted on the left and the larger on the right, except for the first formant region with the 200-Hz fundamental. Here the center frequencies are plotted in reversed order (i.e., 600 Hz, 500 Hz). For the 100-Hz fundamental, the jnd's for the sine-phase, random-phase, and Schroeder-phase relations are shown as square, circular, and hourglass symbols, respectively. For the 200-Hz fundamental, the jnd's for these three phase relations are shown as triangular-up, triangular-down, and umbrella symbols. For clarity, the jnd's for the 100- and 200-Hz fundamentals are connected by a solid and a dashed line, respectively. Only subject JL contributed jnd's for the Schroeder-phase relation. The jnd's for the filtered noise bands are shown as asterisks, connected by a dotted line. For the noise bands the value of the center frequency relative to the position of the harmonics is not relevant; we chose 500 and 600 Hz for the first formant and 2000 and 2100 Hz for the second formant, plotted at the positions /I\ and /II\, respectively.

The average jnd's for the sine-phase and the random-phase relations are shown in Fig. 3. Each row contains average jnd's for the four combinations of formant frequency and slope. The columns on the left and on the right contain averages for the fundamentals of 100 and 200 Hz, respectively. For clarity the jnd's for the sine and the random phase are connected by a solid and a dashed line, respectively. The averages for the filtered noise bands (the asterisks) are connected by a dotted line. For comparison, the predictions of the modified place model (see the Introduction) are shown for each condition by means of the little line segments at the sides of each panel. The settings of this model were: a Q of 5 for the roex filters, and a threshold of 2 dB. With this modified place model much of our data, of both the first and the second experiment, can be explained. The predictions of this model are introduced in Fig. 3 to enable easy distinction of the jnd's that are in agreement with this model from those that require a different explanation. The solid and the dotted line segments represent the expected jnd's for the harmonic stimuli and the noise bands, respectively.

In the region of the first formant ($F_c \approx 500$ Hz), the average jnd's for the filtered noise bands are 1.7% for a slope of 50 dB/oct, and 1.0% for a 100-dB/oct slope. These values imply just-audible level differences in the flanks of the stimuli of 1.2 and 1.4 dB, respectively. For the 100-dB/oct slope, the noise band jnd's for the two center frequencies differ by a factor of 2. A likely explanation for this difference arises from the specific random phases of the noise bands: The perceived pitch of the noise bands is found to vary slightly with the different phase relations, which affects

the jnd's. In the second formant region ($F_c \approx 2000$ Hz) the average jnd's are 0.8% and 0.45% for the slopes of 100 and 200 dB/oct, respectively. These jnd's correspond to level differences in the flanks of the stimuli of 1.2 and 1.3 dB. So, in all cases we find that a level difference of slightly over 1 dB is necessary for discrimination. When comparing the jnd's for the filtered noise bands with the predictions of the modified place model, we see a good correspondence in the three uppermost rows of Fig. 3. For the stimuli with a 200-dB/oct slope in the second formant region, the modified place model predicts somewhat larger jnd's than were measured.

When considering the average jnd's for the harmonic stimuli in the first formant region, we find hardly any influence of the fundamental, the phase relation, and the position of the center frequency relative to the harmonics. The jnd's for the slope of 100 dB/oct are nearly all a factor of 2 smaller than those for corresponding conditions with the 50-dB/oct slope. An exception to these observations is found for the 200-Hz fundamental with the 100-dB/oct slope when the center frequency lies at a signal harmonic. Here we find very large jnd's. Note that this also holds for the individual jnd's in Fig. 2, including those for Schroeder phase of subject JL. These large jnd's can be easily understood in terms of a place model when we consider the spectrum of the harmonic stimuli for this condition (see Fig. 1, panel h). The spectrum of the reference tone contains only one frequency component at 600 Hz. A considerable increase in the center frequency (of about 25 Hz) is needed to make a second component appear at 800 Hz; because of the loudness correction and the roving stimulus level, there is no perceptual change in the stimulus until this second component appears. When comparing the jnd's for the harmonic stimuli of the first formant region with the predictions of the modified place model, we see a good correspondence.

For the second formant region the situation is more complex. For the slope of 100 dB/oct, the phase relations affect the jnd's when the center frequency lies halfway between two harmonics. For the 100-Hz fundamental, a large difference is seen in the average jnd's for the two phase relations, which is also present in the individual results of Fig. 2. For the 200-Hz fundamental, a small difference is seen between the jnd's for the sine and the random phase. Such a difference can also be found in the individual results, though not in the same direction for all subjects. So, even though the average jnd's are near the predictions of the modified place model, we find some influence of the phase relations on the jnd's. When the center frequency coincides with a harmonic, we see a good correspondence between the average jnd's and the predictions of the modified place model for the 100-dB/oct slope. For the slope of 200 dB/oct we find a consistent influence of the position of the center frequency relative to the harmonics: When the center frequency coincides with a harmonic, the jnd's are larger than those for center frequencies halfway between two harmonics. The jnd's roughly double when the fundamental is increased from 100 to 200 Hz, and there is no influence of the phase relations on the jnd's; this seems in good agreement with the concept of a place model, but all jnd's are smaller than the predictions of the model. So, even though no influence of the phase rela-
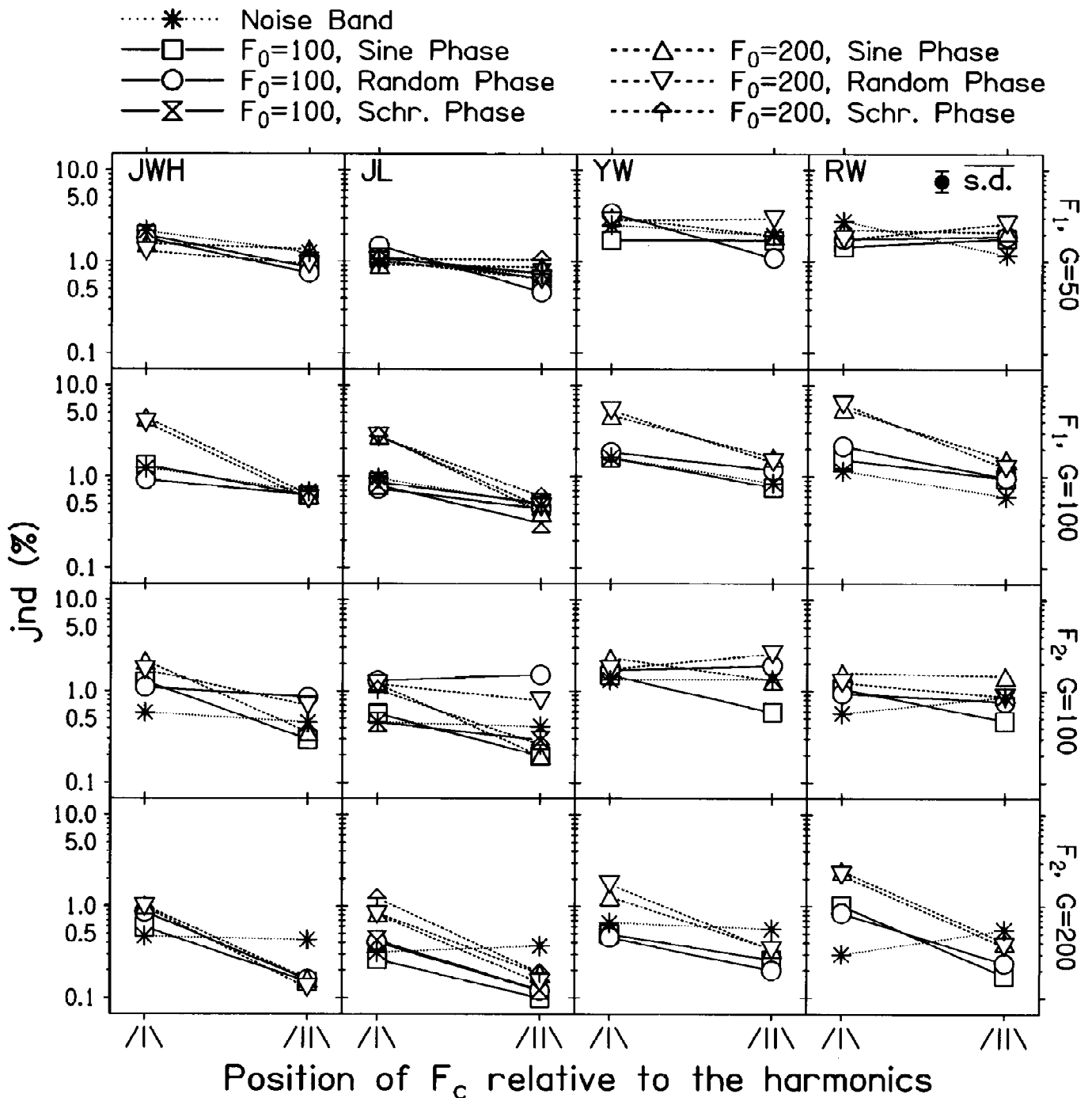
FIG. 2. Individual jnd's for the triangular spectral envelope. Each row contains the jnd's measured for one formant region and slope combination. The error bar in the right top corner indicates the mean standard deviation of the individual results. The jnd's for the 100-Hz fundamental with the sine-phase, the random-phase, and the Schroeder-phase relations are shown as square, circular, and hourglass symbols, respectively. For clarity, these symbols are connected by solid lines. The jnd's for the 200-Hz fundamental, for the same three phase relations, are shown as triangular-up, triangular-down, and umbrella symbols, respectively. These symbols are connected by dashed lines. The jnd's for the noise bands are shown by the asterisks, connected by a dotted line.

tions is seen on the jnd's, temporal effects may have influenced the jnd's.

In summary, all jnd's of the noise bands correspond with level differences of just over 1 dB in the flanks of the stimuli. Under most conditions, they show good correspondence with the modified place model. For the jnd's of the harmonic stimuli in the first formant region, we find no influence of the phase relation, and they are in good correspondence with the modified place model. In the second formant region we find phase effects for the slope of 100 dB/oct when the center

frequency lies between two harmonics. For the 200-dB/oct slope, we find a strong influence of the position of the center frequency relative to the harmonics. Here we find poor correspondence with the modified place model.

In the first and second formant regions (near 0.5 and 2 kHz) the critical bandwidths are approximately 120 and 290 Hz, respectively (Scharf, 1970). So, from the first to the second formant regions of this study, the critical band doubles roughly. A factor of 2 is also present in the stimulus parameters. Therefore in terms of the critical band, several stimuli
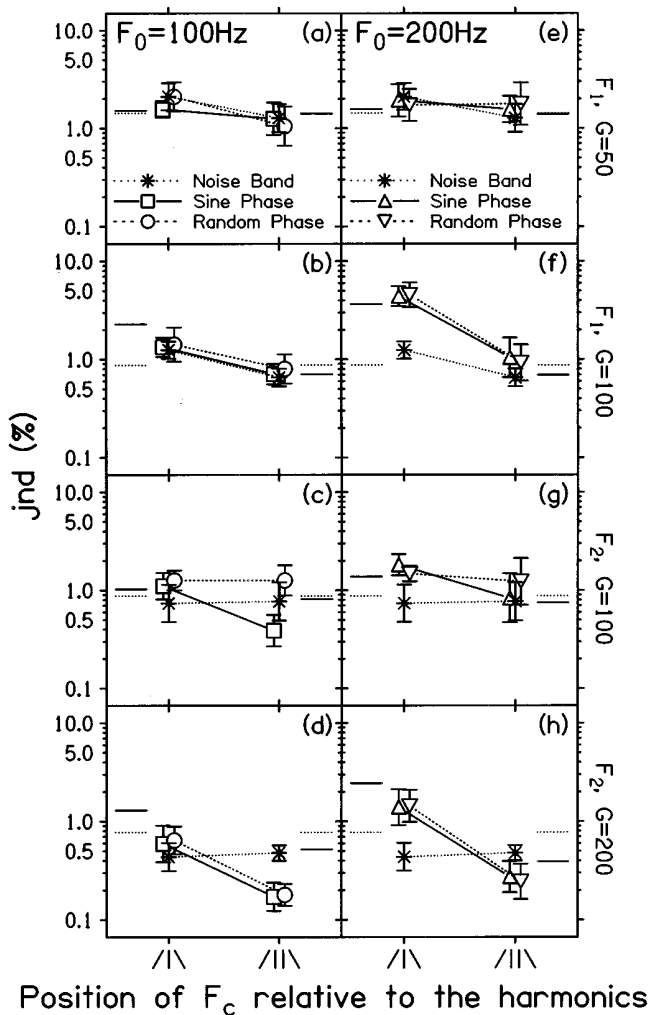
FIG. 3. Average jnd's for the triangular spectral envelope. The rows show jnd's for the same parameters as in Fig. 2. The left column contains the averages for the 100-Hz, and the right column for the 200-Hz fundamental. The meaning of the symbols is the same as in Fig. 2. For clarity, the jnd's for the sine phase are connected by a solid line, and those for the random phase by a dashed line. The error bars indicate the standard deviations of the averages. The little solid and dotted line segments at the side of each panel indicate the predictions of the modified place model for the harmonic stimuli and the noise bands, respectively.

formant region are plotted in panel (a) of Fig. 3, and those for the second formant region in panel (g). These jnd's are similar when the center frequency coincides with a harmonic. When the center frequency lies between two harmonics, the jnd for the sine-phase condition of the second formant region is somewhat smaller than the others. Here we find phase effects in the individual jnd's, indicating that the phase relation influences the discrimination process. For the second combination (with the steeper spectral slopes), the jnd's for the first formant region are plotted in panel (b) of Fig. 3, and those for the second formant region in panel (h). These jnd's are similar when the center frequency coincides with a harmonic. When the center frequency lies between two harmonics, the jnd's for the first formant region are much larger than those for the second formant region (0.7% and 0.25%, respectively). This indicates that for the second formant region temporal aspects of the stimuli are likely to play a role in the discrimination process.

III. EXPERIMENT 2: THE ROUNDED SPECTRAL ENVELOPE

The results of the second experiment are presented in Fig. 4. The format of this figure is the same as that of Fig. 2. Each row contains six panels with individual results. The meaning of the symbols is the same as in Fig. 2. In addition, the jnd's for the Klatt phase for the 100-Hz and the 200-Hz fundamentals are shown as diamond and picnic-table symbols, respectively. All subjects contributed jnd's for two or three conditions in such a way that three subjects participated for the sine, the random, and the Klatt phase, two subjects for the Schroeder phase, and five for the noise bands.

The averages of the results are shown in Fig. 5. This figure has the same format as Fig. 3. For clarity, the jnd's for sine-phase, random-phase, Klatt-phase, and Schroeder-phase relations are connected by a solid, a short dashed, a medium dashed, and a long dashed line, respectively. Note that, due to the distribution of subjects over the various conditions, the averages for each phase relation and for the noise bands are calculated over different groups of subjects. The predictions of the modified place model are shown for harmonic stimuli and for noise bands at the sides of the panels with the little solid and dotted line segments, respectively.

For the first formant region, the average jnd for the noise bands is 1.5% for the slope of 50 dB/oct, and 0.9% for the 100-dB/oct slope. At the steepest parts of the flanks of the stimulus spectra, these jnd's correspond to level differences of 1.1 and 1.3 dB, respectively. For the second formant region, the average jnd for the noise bands is 0.8% for the slope of 100 dB/oct and 0.55% for the 200-dB/oct slope, corresponding to a maximal level difference of 1.2 and 1.6 dB, respectively. The values for both the jnd's and the level differences are similar to those found for the triangular spectral shapes, though larger values would rather be expected for such narrow noise bands because of their amplitude fluctuations. Moreover, when comparing the jnd's for the filtered noise bands with the predictions of the modified place model, we see that the measured jnd's are somewhat smaller than the predictions. This difference is largest for the second

of the first formant region show a rough equivalence to those of the second formant region. First, a factor of two is present in the fundamental of 100 Hz combined with the slope of 50 dB/oct [Fig. 1, panels (a) and (b)] as compared to the 200-Hz fundamental in combination with the 100-dB/oct slope [Fig. 1, panels (m) and (n)]. Second, it is present in the 100-Hz fundamental combined with the 100-dB/oct slope [Fig. 1, panels (c) and (d)] as compared to the 200-Hz fundamental in combination with the 200-dB/oct slope [Fig. 1, panels (o) and (p)]. (For this reason the abscissa scales of the leftmost and rightmost columns of Fig. 1 differ by the same factor of 2, which gives the stimuli of the compared groups the same appearance in this figure.) So, in terms of spectral resolution of the ear, these two combinations of conditions are roughly similar. Therefore a spectral explanation of the discrimination process can only be viable when the jnd's for the harmonic stimuli are roughly similar too. For the first combination (with the shallow spectral slopes), the jnd's for the first

FIG. 4. Individual jnd's for the Klatt spectral envelope. The format of this figure, and the meaning of the symbols are the same as in Fig. 2. The jnd's for the Klatt-phase relation are shown as diamonds for the 100-Hz fundamental, and as picnic-table symbols for the 200-Hz fundamental.

formant region, which implies that the present modified place model may not be applicable to these jnd's.

In the averaged results for the harmonic stimuli in the region of the first formant, we see no influence of the phase relation on the jnd for the 200-Hz fundamental. These jnd's are not affected by the position of the center frequency relative to the harmonics. They are all larger than those for the noise bands. This may be understood from the large spacing of the components: A change in the center frequency of a spectral envelope cannot be transformed optimally into component level changes when no components are present in the steepest parts of its slopes. For the 100-Hz fundamental, we seem to find an influence of the phase relations on the average jnd's. However, this effect is apparently caused by the distribution of the subjects over the various phase relations, since the individual results in Fig. 4 show no phase effects. So, in this case the phase effects shown by the average jnd's are artifacts of the averaging. When we ignore the spread in the jnd's arising in this way, we see a good correspondence between the jnd's for the harmonic stimuli of the first formant region and the predictions of the modified place model.

In the second formant region, we find a clear influence of the phase relation on the average jnd's for the 100-Hz fundamental; this influence is found in the individual results of four of the six subjects. It is systematic: For the sine phase the difference between the jnd's for center frequencies positioned on a harmonic and between two harmonics is larger than for the other phases, and the jnd's for the random phase are usually the largest. For the 200-Hz fundamental no influence of the phase relation on the individual jnd's is found for the slope of 200 dB/oct: The observed spread in the averages is due to the averaging over a different group of subjects for each phase relation. For the 100-dB/oct slope, some influence of the phase relation can be found in the individual jnd's of Fig. 4, though not in the same direction for all subjects. The observed spread in the averages, however, is largely due to the averaging over different subjects for each phase relation. When comparing the jnd's for the harmonic stimuli with the predictions of the modified place model, we see a good correspondence for the 200-Hz fundamental; correspondence is poor for the 100-Hz fundamental.

In summary, all jnd's of the noise bands correspond with a level difference in the flanks of the stimuli of just over 1 dB. Nevertheless, these jnd's are consistently somewhat smaller than the predictions of the modified place model. For the harmonic stimuli, we find no consistent influence of the phase relation on the jnd's for the first formant region in general, and we find small phase effects for the second formant region in the case of the 200-Hz fundamental. Under these conditions, the jnd's are close to the predictions of the modified place model. For the second formant region and the 100-Hz fundamental, we find clear phase effects in the jnd's,
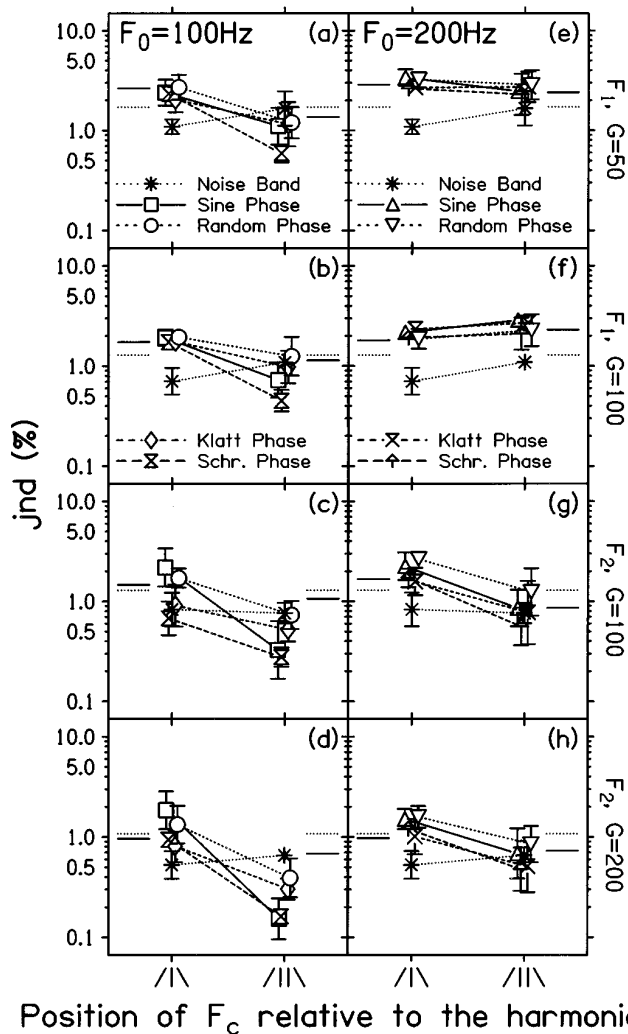
FIG. 5. Average jnd's for the Klatt spectral envelope. The format of this figure is the same as in Fig. 3, the meaning of the symbols is the same as in Fig. 4. The error bars indicate the standard deviations of the averages.

indicating that here the discrimination is influenced by a temporal process.

As for experiment 1, we can compare the results of the first formant region for the 100-Hz fundamental with those of the second formant region for the 200-Hz fundamental. In terms of the critical bandwidth, the stimuli under these two conditions show a rough equivalence. A spectral explanation of the jnd's can only be viable when the jnd's in panels (a) and (b) of Fig. 5 are roughly in agreement with those in panels (g) and (h), respectively. For both combinations, the jnd's for the first formant region are close to the corresponding ones for the second formant region. Furthermore, these jnd's are all in the vicinity of the predictions of the modified place model.

IV. GENERAL DISCUSSION

A. Applicability of the place model

When comparing the jnd's for the harmonic stimuli under the corresponding conditions of both experiments as they are plotted in Figs. 3 and 5, we find for both formant regions that correspondence is best for the combination of the 100-Hz fundamental and the shallow slope. For the combi-

nation of the 200-Hz fundamental and the steep slope correspondence is worst. For the average noise band jnd's, resemblance is good for the corresponding conditions of both experiments. So, we find that correspondence between the jnd's of both experiments is better when the peaks of the spectra resemble each other more closely (see Fig. 1). As was expected, we find smaller jnd's for the steep rather than for the shallow slopes, for the low rather than for the high fundamental, and for center frequencies halfway between harmonics rather than for those coinciding with a harmonic.

For the second formant region, Lyzenga and Horst (1995) suggested a modified place model with a 1-dB excitation-difference threshold to explain their data for the triangular spectral envelope with a center frequency halfway between two harmonics. However, in the present experiments we often find phase effects for similar stimuli, indicating a temporal rather than a spectral discrimination process. Below we will propose a temporal discrimination mechanism that describes the jnd's found for these stimuli.

The model, suggested by Lyzenga and Horst (1995) to explain their jnd's for stimuli with the trapezoidal envelope, is equivalent to an original (i.e., nonmodified) place model with an excitation-difference threshold of 1 dB. For a good correspondence between place model predictions and the present data we need a modified place model with a 2-dB threshold. This threshold represents an excitation difference of 1 dB in both the low- and the high-frequent flanks of the stimuli. (Both models use roex filters with a Q of 5.) For the stimuli with a trapezoidal envelope used by Lyzenga and Horst the predictions of these two models hardly differ; the squared relative error: $\Sigma([\text{jnd}-\text{prediction}]/\text{jnd})^2$ between the data and the model predictions is 5.4 for the original place model, and 7.7 for the modified place model (both summed over 15 jnd's). For the present data the modified place model describes the data considerably better; the squared relative error is 42.1 for the original place model, as compared to 12.3 for the modified place model (both summed over 32 jnd's for experiment 1 and 64 jnd's for experiment 2). When disregarding the above-mentioned jnd's for the second formant region with a center frequency halfway between two harmonics, the squared relative error for the original place model is 36.1, and for the modified place model it is 6.6 (both summed over 24 jnd's for the first and 48 jnd's for the second experiment). So, the use of the modified rather than the original version of the place model greatly improves the accuracy of the modeling.

For the first formant region, we find that the modified place model provides a good explanation of our jnd's. This is in good agreement with the observations that for both experiments the jnd's of the first formant region show no phase effects and that the jnd's for the noise bands correspond with a level difference in the flanks of the stimuli of just over 1 dB. Therefore it is likely that spectral cues dominate the discrimination process here.

In the second formant region for the 100-Hz fundamental, we find phase effects in the jnd's of both experiments. Under these conditions, three stimulus components fall within one critical band (of 250 Hz), so their relative phases can influence the waveform. This is in agreement with find-

ings of Edwards and Viemeister (1994) and of Narayan and Long (submitted). Because of its insensitivity to the phase relations of the stimulus components, the modified place model clearly cannot explain the data for the 100-Hz fundamental. For the 200-Hz fundamental we find good correspondence between the predictions of the modified place model and the data for the stimuli with the Klatt envelope and with the triangular envelope with the 100-dB/oct slope. However, for these conditions, except for the Klatt envelope with center frequencies that coincide with a harmonic, alternative explanations will be given below. The modified place model fails to describe the data for the triangular envelope with the 200-dB/oct slope.

For the noise stimuli of the second formant region, we only find good agreement with the predictions of the modified place model for the triangular envelope with the slope of 100 dB/oct. We have tried various schemes for decreasing the discrepancies between the jnd's for the 200-dB/oct slope and the model predictions, such as using steeper or different filters, using different rules for quantifying the excitation differences (like comparing the threshold with the area under the excitation patterns, rather than with the excitation differences at one or two positions), or introducing a nonlinearity into the model before the filter bank. We found that the discrepancies cannot be reduced without deteriorating the performance of the model for the harmonic stimuli of the first formant region. Yet, all the noise band jnd's correspond with a level difference in the flanks of the stimuli of just over 1 dB. From a spectral point of view, this implies that the auditory filters near the center frequencies of the stimuli are steeper than the slopes of the noise bands. With a Q of 5 the high-frequency slope of the roex filter, that was used in the modified place model, is about 120 dB/oct, which is much smaller than the 200-dB/oct slope of the stimuli. Therefore the present adapted place model fails to describe these data. One possible explanation for the small noise band jnd's for the 200-dB/oct slope arises from the observation that these narrow noise bands start to approximate pure tones with fluctuating amplitudes. This implies that for narrower noise bands the jnd should approach the pure-tone jnd (for roving level experiments), but it can be expected to remain above this due to the frequency jitter present in narrow-band noises. For the slope of 200 dB/oct, our noise band jnd's are between 0.4% and 0.6%. This is in good agreement with Michaels (1957) and Moore (1973) who found a limit of 0.4% for noise bands with decreasing bandwidths, as compared to a constant-level pure-tone jnd of 0.2%. They both proposed a temporal explanation for their results, based on interspike intervals and the integration time of the auditory system in relation to the rate of the amplitude fluctuations.

B. Applicability of temporal models

We related the present jnd's to changes in the amplitude modulation depth of the stimuli. To do this we used a generalization of the concept used by Lyzenga and Horst (1995). We extended their definition of modulation depth to: The decrease in the absolute value of the temporal envelope over one fundamental period, expressed as a percentage of the maximum absolute value of this envelope. The temporal en-
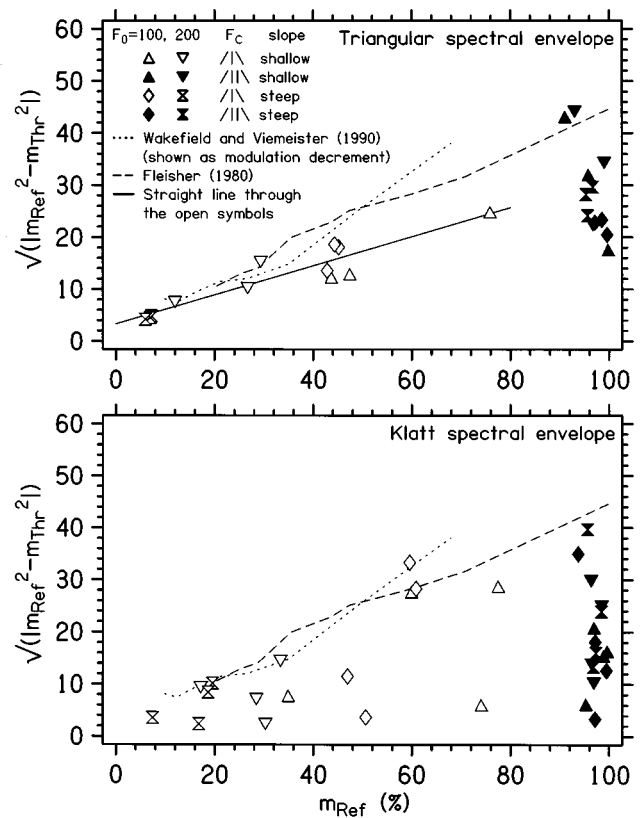
FIG. 6. The difference in modulation power for each reference and its just-noticeable target, as a function of the largest modulation depth. The stimuli were filtered with a roex filter with a Q of 5. The dotted lines display data for amplitude-modulated noise with a modulation frequency of 100 Hz of Wakefield and Viemeister (1990). These data have been transformed to a modulation depth decrement. The dashed lines give data of Fleischer (1980) for an amplitude-modulated pure tone of 1 kHz with a modulation frequency of 50 Hz.

velopes were calculated using the Hilbert transforms of the stimuli. For the stimuli of the second formant region with the triangular envelope and a center frequency coinciding with a harmonic, we found a linear relation between the initial modulation depth and the difference in modulation power between the just-noticeably different stimuli (very similar to the results in the upper panel of Fig. 9 of Lyzenga and Horst, 1995). However, correspondence was not found for the stimuli with the Klatt envelopes. This may be connected with the often very irregular temporal envelopes of the Klatt stimuli (see the left column of Fig. 7). To smooth these temporal envelopes, and to approximate the stimuli and their temporal envelopes as they appear within separate frequency channels of the auditory system, we applied an auditory filter to all the stimuli. We used a roex filter with a Q of 5, as was also used in the modified place model. After auditory filtering, the correspondence between the modulation depth and the jnd was good again for the triangles but still poor for the Klatt stimuli. This is shown in Fig. 6. This figure shows the difference in modulation power (the square root of the difference of the squares of the modulation indexes) between the just-noticeably different stimuli as a function of the maximal modulation depth for the triangular (top panel) and the Klatt envelope (bottom panel). For comparison, data are displayed of Wakefield and Viemeister (1990) for amplitude-

modulated noise and a 2IFC task, and of Fleischer (1980) for amplitude-modulated pure tones and a same–different task.

For the stimuli with a center frequency coinciding with a harmonic (open symbols), the line described by the modulation power difference as a function of the maximal modulation depth for the triangular envelopes coincides reasonably well with the data of Wakefield and Viemeister (1990) and Fleischer (1980). But these data only function as an upper limit for modulation differences found for the Klatt envelope. For these stimuli the differences in modulation power are mostly smaller than for the triangles. So, here we find an essential difference connected with the bandwidth of the stimuli: For the triangular stimuli the modulation depth provides a good explanation of the jnd's found for stimuli with a center frequency coinciding with a harmonic, but for the Klatt stimuli this is not the case. When fitting a straight line to the data for the triangles with center frequencies coinciding with a harmonic (open symbols) using a least-squares method, we find the relation: $y = 0.28*x + 3.3$, shown as the solid line in the upper panel of Fig. 6. All data in the open symbols lie within 35% of this line. Thus the actual model expectations are those center frequency differences that yield modulation depths in accordance with this line.

For the first formant region, we found a correlation between the modulation depth and the jnd only when no auditory filtering was applied in the calculation of the modulation depths of the stimuli. Under these conditions, the correlation between the modulation power and the jnd was comparable to the one found for these stimuli in the second formant region. For the stimuli with the Klatt envelope, the correlation found for the triangles only provided an upper limit for the modulation power. After auditory filtering of the stimuli with a frequency equal to their center frequency, the modulation depths and differences were negligible. For the first formant region the critical bandwidth is about 120 Hz as compared to the 100- and 200-Hz fundamentals. Therefore all the stimulus components are rather well resolved by the auditory system, and modulation depth is not likely to be involved in the discrimination.

In analogy to the definition of amplitude modulation that we used, a definition of the phase or frequency modulation can be formulated for the present stimuli: The absolute shift in the phase over one fundamental period relative to a sinusoid with a frequency equal to the center frequency of the stimulus. When the phase is expressed in radians, this value is two times the frequency modulation index for a sinusoidally modulated pure tone. For both the first and the second formant regions, we did not find any correspondence between the phase or frequency modulation depth and the jnd's.

Alternatives for explaining the jnd's for harmonic stimuli might be found in the EWAIF or the IWAIF fractions, which are defined as:

$$\Delta \text{EWAIF} = \frac{\int_0^T E_r(t) f_r(t) dt}{\int_0^T E_r(t) dt} - \frac{\int_0^T E_t(t) f_t(t) dt}{\int_0^T E_t(t) dt}, \qquad (3)$$

$$\Delta \text{IWAIF} = \frac{\int_0^T E_r^2(t) f_r(t) dt}{\int_0^T E_r^2(t) dt} - \frac{\int_0^T E_t^2(t) f_t(t) dt}{\int_0^T E_t^2(t) dt}, \qquad (4)$$

where T is the period of the stimuli, and $E_r(t)$ and $E_t(t)$ are the temporal envelopes for the reference and the just-noticeably different target, respectively; while $f_r(t)$ and $f_t(t)$ are the corresponding instantaneous frequencies. These instantaneous frequencies were calculated from the first derivative of the instantaneous phases generated by the Hilbert transforms of the stimuli. The envelope-weighted averaged instantaneous frequency was introduced by Feth (1974) to describe the pitch of unresolved two-tone complexes. Many of our stimuli of the second formant region are unresolved, and a number of them resemble two-tone complexes. If the EWAIF or IWAIF fractions provide an explanation of the jnd's, we expect a constant difference for just-noticeably different stimuli. We calculate these frequency differences both for the stimuli filtered with an auditory filter at their center frequency, and for the unfiltered stimuli. For the auditory filtering, we used a roex filter with a Q of 5. For both the filtered and the unfiltered stimuli of the second formant region, the EWAIF and IWAIF fractions showed little correspondence with the jnd's. So, neither EWAIF nor IWAIF provides a viable explanation for our data.

For the stimuli in the second formant region with a center frequency halfway between two harmonics (the closed symbols in both panels of Fig. 6) neither the modified place model nor the modulation depth provide an explanation of the data. Under these conditions we often find a sudden change in the direction of the temporal envelope of the triangular stimuli. These changes in direction occur when the temporal envelope touches zero, where it shows a sharp notch. Such notches produce peaks in the second-order derivative of the temporal envelope, the size of which give a measure of the sharpness of the notch. This is illustrated in the top half of Fig. 7 for triangular stimuli with a fundamental of 100 Hz and a slope of 200 dB/oct. We see that the peaks in the second-order derivative depend strongly on the center frequency, and that they are influenced in their size and shape by the phase relation of the components.

The temporal envelopes of the stimuli with the Klatt envelope differ strongly from those with the triangular envelope; the first behave very erratically, as do their second-order derivatives (see the bottom half of Fig. 7). The differences between the two types of stimuli are much reduced when both are filtered with a simulated auditory filter: we used a roex filter with a Q of 5. We have tried a number of phase relations for the auditory filter: zero phase, an arctangent phase, and (inspired by Kohlrausch and Sander, 1995) an exponential phase with various amplitudes for the phase sweep. The correspondence between the temporal model, as described in the next paragraph, and the data was best for the exponential-phase relation shown in Fig. 8.

In the temporal model, the second-order derivative of the temporal envelope of the filtered stimuli is used to predict jnd's for the second formant region. When a peak in the second-order derivative of the reference tone is present, the center frequency is increased until this peak is reduced to a certain threshold value. Optimal correspondence between model and data was found for a threshold equal to one-third of the initial peak value. When the threshold is reached within a reasonable center frequency shift (64 Hz), the cor-
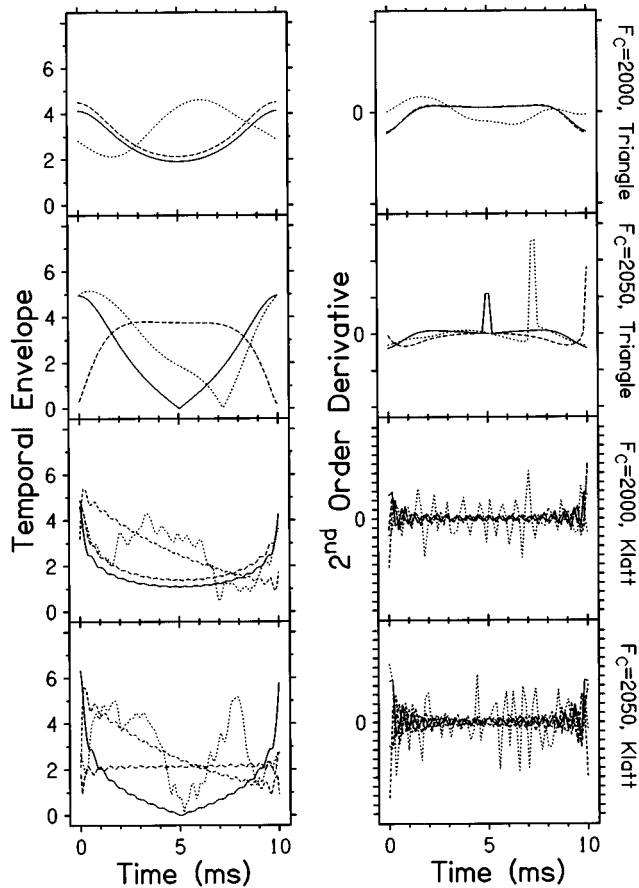
FIG. 7. Temporal envelopes (left column) and their second-order derivatives (right column) shown for the triangular spectral envelope in the upper two rows, and for the Klatt envelope in the lower two rows. Each panel shows curves for sine phase (solid lines), Schroeder phase (long-dashed lines), and random phase (dotted lines). For the Klatt envelope, the Klatt-phase relation is shown with the short-dashed lines. The center frequency is indicated on the right. For the triangular envelope with the center of frequency of 2050 Hz, we find peaks in the second-order derivatives for all three displayed phase relations.

FIG. 8. The amplitude and phase characteristic of the assumed auditory filter.

This temporal mechanism is analogous to a model proposed by Green *et al.* (1992) for spectral shape discrimination of narrow-band sounds. Their model uses changes in the power spectrum of the temporal envelope as a discrimination cue, whereas our model uses changes in the sharpness of the minima in the temporal envelope. These two properties are related: A sharp change in the temporal envelope will produce one or more peaks in the high-frequency region of the power spectrum of the temporal envelope. When changing

responding center frequency difference is taken as the expected jnd. This procedure works only for stimuli with a center frequency halfway between two harmonics. When the center frequency coincides with a harmonic, either no peak is present, or it does not sufficiently change its size. In Fig. 9, the predictions of this temporal model are compared with the jnd's of the second formant region for both experiments. In this figure only jnd's under the conditions with center frequencies halfway between two harmonics are considered. For the triangular envelope, the jnd's plotted for the Schroeder phase are the individual data of subject JL. For the random phase, under the condition of the Klatt envelope with the 100-Hz fundamental and the 100-dB/oct slope, the model failed to produce an expected value. However, under this condition we find a good correspondence with the modified place model (see Fig. 5). In general we see a good correspondence between temporal model and data in Fig. 9. The presence of phase effects in these data indicates that a temporal mechanism is used in the discrimination process. The correspondence between the predictions of the present model and the data indicates that such a temporal mechanism may well be based on changes in the temporal envelopes of the stimuli.

FIG. 9. The average jnd's, for the second formant region and for center frequencies halfway between harmonics, compared with the results of the temporal model. The averages are replotted in the center of the panels, and the predictions of the model are plotted at the sides, using the same symbols for the phase relations as in Fig. 5. The column on the left displays this comparison for the triangular envelope, and the column on the right for the Klatt envelope. The slope is 100 dB/oct for the top row, and 200 dB/oct for the bottom row.

the center frequency of our stimuli, sharp changes in the temporal envelope may either appear or disappear, and so will the corresponding peaks in the power spectrum of the temporal envelope. This is comparable to the situations displayed in the bottom panels of Figs. 3 and 4 of Green *et al.* (1992), where peaks in the power spectrum are seen in the presence of a 10-dB signal, whereas they are absent without the signal. The calculation of the power spectrum of the temporal envelope of a tone requires an algorithm considerably more complex than ours. Still, our method produced excellent agreement with the present data.

V. CONCLUSIONS

The behavior of the jnd's of the noise bands is roughly in agreement with the predictions of the modified place model. For very narrow bands they approach the roving-level pure-tone jnd. The jnd's for the harmonic stimuli of the first formant region are in good agreement with the modified place model. For the harmonic stimuli of the second formant region, the explanation depends on the shape of the spectral envelope of the stimuli and on the position of the center frequency relative to the harmonics. Here we find three mechanisms involved in the discrimination process, where, under each stimulus condition, the best performing mechanism determines the jnd. When the center frequency coincides with a harmonic, the jnd's for the Klatt envelope are in agreement with the modified place model, and those for the triangular envelope with modulation depth differences. For the stimuli with a center frequency halfway between two harmonics, we find good correspondence between the jnd's and a temporal model that is based on the sharpness or spikiness of the minima in the temporal envelope (represented by the peaks in the second-order derivative of the temporal envelope).

ACKNOWLEDGMENTS

Bilsen, F. A. (**1973**). ''On the influence of the number and phase of harmonics on the perceptibility of the pitch of complex signals,'' Acustica **28,** 60–65.

Duifhuis, H. (**1970**). ''Audibility of high harmonics in a periodic pulse,'' J. Acoust. Soc. Am. **48,** 888–893.

Duifhuis, H. (**1972**). ''Perceptual analysis of sound,'' Ph.D. dissertation, Eindhoven University of Technology, Eindhoven, The Netherlands.

Edwards, B. W., and Viemeister, N. F. (**1994**). ''Modulation detection and discrimination with three-component signals,'' J. Acoust. Soc. Am. **95,** 2202–2212.

Feth, L. L. (**1974**). ''Frequency discrimination of complex periodic tones,'' Percept. Psychophys. **15,** 375–379.

Fleischer, von H. (**1980**). ''Subjektive Grösse von Unterschieden im Amplituden-modulationsgrad von Sinustönen,'' Acustica **46,** 31–37.

Fletcher, H. (**1940**). ''Auditory patterns,'' Rev. Mod. Phys. **12,** 47–65.

Gagné, J., and Zurek, P. M. (**1988**). ''Resonance-frequency discrimination,'' J. Acoust. Soc. Am. **83,** 2293–2299.

Goldstein, J. L., and Srulovicz, P. (**1977**). ''Auditory nerve spike intervals as an adequate basis for aural frequency measurements,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 337–346.

Green, D. M., Berg, B. G., Dai, H., Eddins, D. A., Onsan, Z. A., and Nguyen, Q. T. (**1992**). ''Spectral shape discrimination of narrow-band sounds,'' J. Acoust. Soc. Am. **92,** 2586–2597.

Henning, G. B. (**1966**). ''Frequency discrimination of random-amplitude tones,'' J. Acoust. Soc. Am. **39,** 336–339.

Hoekstra A., and Ritsma, R. J. (**1977**). ''Perceptive hearing loss and frequency selectivity,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 263–271.

Horst, J. W., Javel, E., and Farley, G. R. (**1990**). ''Coding of spectral fine structure in the auditory nerve. II. Level-dependent nonlinear responses,'' J. Acoust. Soc. Am. **88,** 2656–2681.

Horst, J. W., Javel, E., and Farley, G. R. (**1992**). ''Coding of fundamental frequency in auditory nerve fibers: Effects of signal level and phase spectrum,'' in *The Auditory Processing of Speech*, edited by M. E. H. Schouten (Mouton de Gruyter, Berlin), pp. 29–36.

Horst, J. W., Ritsma, R. J., and Wit, H. P. (**1984**). ''Frequency discrimination in quiet and noise for signals with triangular spectral envelopes,'' J. Acoust. Soc. Am. **76,** 1067–1075.

Klatt, D. H. (**1980**). ''Software for a cascade/parallel formant synthesizer,'' J. Acoust. Soc. Am. **67,** 971–995.

Kohlrausch, A., and Sander, A. (**1995**). ''Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets,'' J. Acoust. Soc. Am. **97,** 1817–1829.

Lundeen, C., and Small, A. M. (**1984**). ''The influence of temporal cues on the strength of periodicity pitches,'' J. Acoust. Soc. Am. **75,** 1578–1587.

Lyzenga, J., and Horst, J. W. (**1995**). ''Frequency discrimination of band-limited harmonic complexes related to vowel formants,'' J. Acoust. Soc. Am. **98,** 1943–1955.

Michaels, R. M. (**1957**). ''Frequency difference limens for narrow bands of noise,'' J. Acoust. Soc. Am. **29,** 520–522.

Moore, B. C. J. (**1973**). ''Frequency difference limens for narrow bands of noise,'' J. Acoust. Soc. Am. **54,** 888–896.

Moore, B. C. J. (**1977**). ''Effects of relative phase of the components on the pitch of three-component complex tones,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 349–362.

Moore, B. C. J., and Peters, R. W. (**1992**). ''Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity,'' J. Acoust. Soc. Am. **91,** 2881–2893.

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,'' Hearing Res. **28,** 209–225.

Narayan, S. S., and Long, G. R. (**submitted**). ''Monaural phase sensitivity: Effects of frequency resolution and stimulus level,'' J. Acoust. Soc. Am. (submitted).

Patterson, R. D. (**1973**). ''The effects of relative phase and the number of components on residue pitch,'' J. Acoust. Soc. Am. **53,** 1565–1572.

Patterson, R. D., 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), pp. 123–177.

Plomp, R., and Steeneken, H. J. M. (**1969**). ''Effect of phase on the timbre of complex tones,'' J. Acoust. Soc. Am. **46,** 409–421.

Ritsma, R. J., and Engel, F. L. (**1964**). ''Pitch of frequency-modulated signals,'' J. Acoust. Soc. Am. **36,** 1637–1644.

Scharf, B. (**1970**). ''Critical bands,'' in *Foundation of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol. 1, pp. 159–202.

Schroeder, M. R. (**1970**). ''Synthesis of low-peak-factor signals and binary sequences with low autocorrelation,'' IEEE Trans. Inf. Theory **16,** 85–89.

Stevens, K. N. (**1952**). ''Frequency discrimination for damped waves,'' J. Acoust. Soc. Am. **24,** 76–79.

Wakefield, H. W., and Viemeister, N. F. (**1990**). ''Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise,'' J. Acoust. Soc. Am. **88,** 1367–1373.

Zwicker, E. (**1970**). ''Masking and psychological excitation as consequences of the ear's frequency analysis,'' in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden), pp. 376–396.

Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners

Marina M. Rose and Brian C. J. Moore

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

This study examined the perceptual grouping of rapid tone sequences for listeners with normal hearing and listeners with unilateral and bilateral cochlear hearing loss. The sequence ABA–ABA– was used, where A and B represent sinusoidal tones bursts (10-ms rise/fall, 80-ms steady state, 20-ms interval between tones) and – represents a silent interval of 120 ms. Tone A was fixed in frequency at 250, 500, 1000, or 2000 Hz. Tone B started with a frequency well above or below that of tone A, and its frequency was swept towards that of tone A so that the frequency separation between them decreased in an exponential manner. Listeners were required to indicate when they could no longer hear the tones A and B as two separate streams, but heard only a single stream with a ''gallop'' rhythm. This is called the fission boundary. For the normally hearing listeners, the separation between tones A and B at the fission boundary was roughly independent of the frequency of tone A when expressed as the difference in number of ERBs (delta E) between A and B, which is consistent with a recent model of stream segregation [M. W. Beauvois and R. Meddis, J. Acoust. Soc. Am. **99**, 2270–2280 (1996)]. For the unilaterally hearing-impaired listeners, there was no consistent difference in the delta E magnitudes across ears, even though the auditory filters were broader in the impaired ears. This is not consistent with the theory of Beauvois and Meddis. The bilaterally hearing-impaired listeners sometimes showed delta E magnitudes within the normal range, and sometimes showed larger than normal delta E magnitudes. The results are discussed in terms of the factors that might influence perceptual stream formation in hearing-impaired listeners. © *1997 Acoustical Society of America.* [S0001-4966(97)04409-3]

PACS numbers: 43.66.Lj, 43.66.Mk, 43.66.Sr [JWH]

INTRODUCTION

People with cochlear hearing impairment often complain of difficulty in following a conversation when background sounds are present. This difficulty may be partly due to reduced audibility; part of the sound may fall below the absolute threshold. However, some difficulty persists even when sounds are amplified so as to ensure audibility (Moore, 1995). This remaining effect has sometimes been explained as a consequence of the impaired frequency selectivity which is found in most people with cochlear damage (Pick *et al.*, 1977; Glasberg and Moore, 1986; Moore, 1995). The broader auditory filters in listeners with a cochlear hearing loss lead to an increased susceptibility to masking by interfering sounds and a reduced ability to discriminate changes in spectral shape. This ability is important for identifying speech sounds such as vowels.

In the literature, discussions of reduced frequency selectivity have mainly concentrated on its effects on simultaneously presented sounds (Moore, 1995). However, it is possible that reduced frequency selectivity also has effects on the perceptual organization of sounds. When several sound sources are present, the auditory system is faced with the task of deciding which elements ''belong'' together, i.e., emanated from the same source. The elements must be put into groups, each corresponding to a particular sound source. This process is known as ''perceptual grouping'' or ''auditory scene analysis'' (Bregman, 1990; Darwin and Carlyon,

1995) and it may be adversely affected by reduced frequency selectivity for reasons discussed below.

When we listen to rapid sequences of sounds, such as those that occur in speech or music, sounds with components falling in similar frequency ranges tend to be grouped together (i.e., perceived as if they come from a single source—called fusion or coherence), whereas sounds with components in different frequency ranges tend to be perceived as different streams (as coming from more than one source—called fission or segregation) (van Noorden, 1975; Bregman, 1990). This principle is used in polyphonic music, where large musical intervals between successive tones in rapid tone sequences played with only one instrument give rise to the impression of hearing multiple instruments or melodic sequences.

For intermediate frequency separations between successive tones in a rapid sequence, the percept may be ambiguous; either fission or fusion may be heard, depending on the attentional set of the subject and the instructions given. If the frequency separation is larger than a critical value, called the temporal coherence boundary (van Noorden, 1975; Beauvois and Meddis, 1996), two streams are always heard. If the frequency separation is less than a (different) critical value, called the fission boundary (van Noorden, 1975; Beauvois and Meddis, 1996), then a single stream is heard. The fission boundary is affected only slightly by the repetition time of the tones, while the temporal coherence boundary increases

markedly with increasing tone repetition time (van Noorden, 1971, 1975).

It has been proposed (Beauvois and Meddis, 1996) that the perceptual organization of such sequences of sounds depends upon the filtering that takes place in the cochlea. The theory of Beauvois and Meddis assumes that the percept of hearing a single stream as opposed to hearing more than one stream depends upon the overlap of the excitation patterns evoked by the sounds in the cochlea; a large degree of overlap leads to fusion while a small degree of overlap leads to fission. The overlap of the excitation patterns depends in turn on the sharpness of the auditory filters (Moore and Glasberg, 1983; Glasberg and Moore, 1990); for a given frequency separation of two successive sounds, the sharper the auditory filters, the smaller the overlap of the excitation patterns will be.

If the theory of Beauvois and Meddis (1996) is correct, then the reduced frequency selectivity that is usually associated with cochlear hearing loss should lead to dramatic changes in the perception of sequences of sounds. For example, in a sequence of tones alternating between two frequencies, a much greater frequency separation between the two tones should be required for the perception of fission as opposed to fusion. One of the main objectives of the experiment reported here was to test if this prediction is correct.

Another prediction of the model of Beauvois and Meddis derives from the concept of a frequency-to-place conversion in the auditory system. The position of the basilar membrane which is excited most by a given frequency varies systematically with frequency; each frequency has its "best place" where it produces maximum excitation. The peripheral auditory system can be considered as an array of overlapping band-pass filters (Fletcher, 1940), where each filter corresponds to a specific place on the basilar membrane.

There is evidence that, for normal hearing, each equivalent rectangular bandwidth (ERB) of the auditory filter corresponds to a constant distance of about 0.89 mm along the basilar membrane (Greenwood, 1961, 1990; Moore, 1986). Glasberg and Moore (1990) proposed a function for converting from a frequency scale to a scale of ERB units. This scale is thought to correspond better to the internal representation of sound than a simple linear or logarithmic frequency scale. The function is

$$E = 21.4 \log_{10}(4.37F + 1), \tag{1}$$

where E is the number of ERBs and F is the center frequency in kHz. For example, the value of the ERB for a center frequency of 1 kHz is about 130 Hz, so an increase in frequency from 935 to 1065 Hz represents a step of one ERB. This scale is conceptually similar to the Bark scale proposed by Zwicker (1961), although it differs somewhat in numerical values.

The theory of Beauvois and Meddis (1996) predicts that the frequency ratio between successive tones at the fission boundary (or at the temporal coherence boundary) should be constant when expressed as the difference in E value between the two tones, delta E, regardless of the center frequency of the tones. This prediction follows from the fact that the overlap of the excitation patterns of the two tones is

roughly constant when delta E is held constant. Beauvois and Meddis assessed this prediction using the data of Miller and Heise (1950), who examined the perception of a tone sequence alternating between two frequencies with a rate of ten tones per second. Miller and Heise determined the musical interval at which "the trill seems to break." They named this tone interval the "trill threshold." It is not clear from their description whether they measured the temporal coherence boundary or the fission boundary; indeed different listeners may have adopted different criteria. They found that the trill threshold was between 2 and 3 semitones (a frequency ratio between 1.12 and 1.19) over a wide frequency range. When these intervals are converted to delta E values, they are reasonably constant across frequency, but not exactly so; their delta E values range from about 0.5 up to 1.2 over a frequency range from 125–7000 Hz. The trill thresholds are reasonably close to the fission boundaries predicted by the model of Beauvois and Meddis. However, considerable individual variability is apparent in the data and no information is given about within-subject reproducibility.

The experiment reported here was intended to provide a further test of the prediction of the model of Beauvois and Meddis that values of delta E at the fission boundary should be constant, but using a method with well-defined instructions, that was determined in pilot experiments to give stable and reproducible results.

I. METHOD

A. Listeners and choice of test level

Three groups of listeners were tested: normally hearing, unilaterally hearing impaired, and bilaterally hearing impaired. Absolute thresholds are presented in the Results section. The stimuli were presented to the normal ears at a level 40 dB above the absolute threshold at the test frequency (40 dB SL). The level for the impaired ears was chosen so as to give a similar loudness as for the normal ears. These levels are also given in the Results section.

The 12 listeners with normal hearing were aged 21–76 years. One was the author MMR. Three listeners, RPB, RJS, and JEC, aged 48, 67, and 76 years, respectively, had raised thresholds in both ears at high frequencies (presbyacusis). For the frequencies tested in the experiment their hearing was normal or nearly normal.

The four unilaterally hearing-impaired listeners (aged 46, 58, 68, and 80 years) were all diagnosed as having a cochlear loss in their impaired ears. They were all experienced listeners and had previously taken part in a loudness matching study, where they were asked to adjust the level of a tone in one ear to match the perceived loudness of a fixed level tone at the same frequency in the other ear (Moore *et al.*, 1996, 1997b), for various levels of the fixed tone. A straight line was fitted to these results to give the levels required for equal loudness in the two ears; for example, a level of 75 dB might be required in an impaired ear to give the same loudness as a level of 50 dB in the normal ear. With this information, it was possible to present the stimuli to the two ears at a similar loudness level, by calculating the level (dB SPL) required to give 40 dB SL for the normal ear, and

then determining the level required to give the same loudness in the impaired ear. Subject AFW complained that the stimuli presented to the impaired ear with $F_{\text{fix}} = 250$ Hz, at the level calculated by this procedure, were ''too quiet.'' In this one case, the level was increased until he found the sounds to be comfortably loud. This discrepancy may have occurred because his hearing loss fluctuates somewhat over time, and the loudness-matching data were obtained some time before the present study was conducted.

Ten bilaterally hearing-impaired listeners (aged 34–79 years) were used, seven of whom had a moderate flat hearing loss. The others had mild and/or sloping losses. The required testing levels were estimated from their absolute thresholds assuming that they had loudness recruitment, and that loudness reached ''normal'' values at about 100 dB HL; it was assumed that the loudness-recruitment function (dB versus dB) was linear.

B. Stimulus generation

Except for conventional audiometry, all stimuli were produced with a Tucker-Davis System II, using a TDT DD1 16-bit digital-to-analog converter (50-kHz sampling rate). Stimulus level was controlled using a TDT PA4 programmable attenuator, and stimuli were then passed to a head-phone buffer (TDT HB6), a final manual attenuator (Hatfield 2125) and one earpiece of a Sennheiser HD414 headphone. Sound levels are specified as levels close to the eardrum. The calibration was based on average measurements for a group of people, using a Rastronics Portarem 2000 system with a probe microphone within 4 mm of the eardrum. Listeners were tested in an IAC double-walled sound-attenuating chamber.

C. Estimation of absolute thresholds

Absolute thresholds were determined both by conventional audiometry and using a two-alternative forced-choice three-down/one-up staircase method that estimates the 79.4% point on the psychometric function (Levitt, 1971). For the latter, the initial level of the signal was well above the threshold estimated from the previously obtained audiogram. The signal level was initially decreased by 5 dB after three successive correct responses, and was increased by 5 dB after one incorrect response. Correct answer feedback was provided by lights on the response box. The transition from decreasing to increasing level, and vice versa, defined a turn point. After the first four turn points the step size was reduced to 2 dB. A run was terminated after a total of 12 turn points and the threshold was defined as the mean of the signal levels at the last 8 turn points. A run was only accepted if the standard deviation of this mean was smaller than 4 dB. Each threshold reported is the mean of three such runs. If the standard deviation of the mean of the three runs was greater than 3 dB, further estimates were obtained, until the standard deviation of the means was smaller than 3 dB.

The absolute thresholds were measured for sinusoids of frequency 250, 500, 1000, and 2000 Hz, with a steady-state duration of 200 ms and raised-cosine ramps of 10 ms. Observation intervals were separated by 500 ms.

D. Measurement of the fission boundary

The method used here is based on a study of van Noorden (1975). He developed stimuli which were very effective for demonstrating and measuring the difference between fission and coherence, namely the sequence obtained from the sequence ABAB... by omission of every other tone B, as in ABA–ABA–.

In this sequence, the listener can perceive the single stream ABA–ABA–..., which has a characteristic gallop-like rhythm, when the frequency ratio of A and B is small. Alternatively, when the frequency ratio is large, two streams may be heard: A–A... and B–B... . In this case, stream A has twice the tempo of stream B. Fission is always heard when the frequency ratio is large, and coherence is heard when the frequency ratio is small. However, there is an intermediate range where a choice is possible. Here, direction of the listener's attention by instruction can influence the boundary between two percepts. The listener can be instructed to listen selectively, that is, to try to hear the stream A–A–A–... and B–B–B–... separately. Alternatively, the listener can be instructed to listen comprehensively, trying to hear the tones together in a single stream ABA–ABA–... . If the frequency ratio between the two tones is adjusted to find the boundary between the two percepts, then different boundaries are obtained depending on the instructions—''try to hear fission'' or ''try to hear coherence.'' Generally, the former leads to more stable results, and results that are affected less by the exact timing of the tone sequence (van Noorden, 1975). Hence, in the present study, listeners were instructed to try to hear fission for as long as possible when the frequency ratio between successive tones was progressively decreased from a large initial value.

A frequency sweep procedure was used to measure the fission boundary, that is the point where the two tones in the sequence were so close together in frequency that they could not be perceived as two streams any more, but only as one stream with a prominent ''gallop'' rhythm. The frequency of tone A was kept constant (F_{fix}) and the frequency of tone B (F_{var}) was variable. At the beginning of each sweep, tone B started at a frequency far from that of tone A, and progressed towards the frequency of tone A in an exponential manner, described by

$$F_{\text{var}'} = F_{\text{fix}} \left(F_{\text{var}} / F_{\text{fix}} \right)^{(1/s)}, \qquad (2)$$

where $F_{\text{var}'}$ = frequency of tone B in the $(i+1)$th ABA– group, F_{fix} = fixed frequency of tone A, F_{var} = frequency of tone B in the ith ABA–group, and $s = 1.1$ (to give a reasonable steepness for the exponential function).

With this formula, the size of change for the variable frequency depended on the ratio between the variable and the fixed frequency. At the beginning of each sweep, when the interval between tones A and B was large, the interval between tone B in the ith ABA– group and tone B in the $(i+1)$th ABA– group was large, too. When the interval between A and B became smaller, the interval between tone B in the ith and tone B in the $(i+1)$th ABA– group became smaller, too. This was done to determine the boundary between coherence and fission more accurately. After several repeats of this sequence, the tone interval between A and B

	Frequency of tone A (Hz)			
Condition	250	500	1000	2000
I	50–70	50–150	100–200	200–400
II	750–1000	1500–2000	2000–3000	4000–4400

eventually became so small that the listener could not hear two streams any more; only the prominent gallop-like rhythm was perceived. At this point, the listener was asked to press the response button (see instructions), indicating that the fission boundary was reached. Depending on the exact starting value for the frequency of tone B and the listener's perception, this would take about 20 to 40 ABA groups or 10 to 20 s.

Each tone had an 80-ms steady-state duration and 10-ms raised-cosine ramps. Between tones there was a 20-ms silence, and after one ABA group there was a 120-ms pause (empty interval, indicated as –). The initial frequency of tone B, F_{init} was chosen randomly from a previously defined starting range, which depended on F_{fix}. F_{init} started either below (condition I) or above (condition II) F_{fix}, which was either 250, 500, 1000, or 2000 Hz. The range of F_{init} for each condition and value of F_{fix} is given in Table I. The stimuli were presented monaurally, and the two ears were tested alternately. For the unilaterally hearing-impaired listeners, the normal ear was always tested first, and after a short break (about 30 s), the impaired ear was tested at the same fixed frequency.

One run consisted of five sweeps which all had the same fixed frequency for tone A but a randomized starting frequency for tone B. Only the last four out of five sweeps contributed to the mean of one run. Usually the standard deviation of the mean frequency for one run was below 3% of the arithmetic mean value for that run. For some rare occasions, when the standard deviation exceeded 5% of the arithmetic mean value, the run was discarded and a new run

obtained. After finishing one run, the same procedure was repeated with the other ear. One block consisted of eight runs (one run in each ear for each fixed frequency and for one of the two conditions I or II), and would last about 15 min. After one block was completed, the listener had a rest, and then continued with another block of eight runs, either for the same or for the other condition. The data for four blocks for each condition were collected in at least two different experimental sessions, and the overall mean was calculated from the means of the last three blocks. Again, the standard deviations for the overall mean for each listener were usually small, with values that were usually less than 3% and always less than 5% of the arithmetic mean.

E. Instructions

All listeners were instructed to listen selectively. A typical oral instruction was as follows: ''You will hear two alternating tones in one ear. One tone is fixed in frequency, and the other one is varying. Please try to follow the varying one as long as you hear it as a separate pattern. When you cannot follow it separately any more, you will hear a distinct gallop-like rhythm. At this point, press the stop button. As soon as you release the stop button, a very similar sweep will start again. After five such sweeps, there will be a short rest, and the same procedure will be repeated with your other ear. Do you have any questions?''

II. RESULTS

A. Absolute thresholds

Tables II–IV show the absolute thresholds in dB SPL for the individual listeners, as measured with the two-alternative forced-choice method. Table II also shows correction factors for transforming from dB SPL to HL, based on normal absolute thresholds for monaural listening. These were derived from the recent ISO standard for free-field listening (ISO 389-7, 1996), with an allowance for the free field to eardrum transformation (Shaw, 1974) and for the difference between monaural and binaural listening; see

TABLE II. Absolute thresholds in dB SPL for the normally hearing listeners. The age for each listener is given in parentheses. The figure below each frequency gives the number of dB to subtract from the absolute threshold in dB SPL (specified at the eardrum) to get the equivalent hearing level (dB HL).

	Frequency (Hz)							
	250		500		1000		2000	
SPL to HL correction	14		7		6		12	
Ear (r=right, l=left)	r	l	r	l	r	l	r	l
MMJ (21)	10	9	11	8	4	4	12	12
YVJ (21)	7	15	7	1	10	14	9	14
CAO (22)	18	10	13	6	11	0	21	8
YDT (23)	18	17	12	11	4	10	13	12
DAV (26)	25	19	17	10	12	13	24	22
MMR (27)	11	12	7	12	4	8	12	17
IIH (29)	13	15	7	7	10	12	13	25
JIA (34)	14	13	7	9	7	10	13	18
HHA (39)	14	22	12	21	20	16	21	21
RPB (48)	15	11	7	8	7	9	13	10
RJS (67)	15	20	12	14	10	16	25	21
JEC (76)	27	30	22	26	33	29	28	44

TABLE III. As in Table II but for the unilaterally hearing-impaired listeners. The number to the left of each slash (/) is the absolute threshold and the number to the right indicates the level at which testing was performed for each listener at that specific frequency.

| | Frequency (Hz) | | | | | | | |
| | 250 | | 500 | | 1000 | | 2000 | |
Ear (r=right, l=left)	r	l	r	l	r	l	r	l
DDF (45)	83/95	16/56	79/96	18/58	75/93	13/53	73/91	22/62
VVF (57)	74/90	27/67	70/86	20/60	69/85	16/56	70/88	35/75
BOS (67)	25/65	70/80	22/62	44/69	17/57	63/79	25/65	73/91
AFW (80)	69/105	19/59	82/92	15/55	73/88	14/54	70/88	35/75

Moore *et al.* (1997a) for details. There was a good correspondence between thresholds in dB HL derived in this way and those measured by conventional audiometry. However, since the forced-choice measures are less affected by response criterion, they are used in the remainder of this paper.

B. Fission boundary

As described in the Introduction, the theory of Beauvois and Meddis (1996) predicts that the frequency difference between the A and B tones at the fission boundary should be constant when expressed as delta E, the difference in number of ERBs for each frequency. To test this prediction, the results were expressed in terms of delta E, defined as the E value for tone B minus the E value for tone A; Eq. (1) in the Introduction describes the conversion from frequency to E value. Note that throughout this paper the delta E values are derived from the scale proposed for normally hearing listeners (Glasberg and Moore, 1990); the delta E values are *not* based on estimates of the auditory filter bandwidths for the individual listeners.

The four frequencies used for tone $A(F_{\text{fix}})$, namely 250, 500, 1000, and 2000 Hz, have E-values of 6.9, 10.8, 15.6, and 21.2, respectively. Delta E values are negative for condition I and positive for condition II.

1. Normally hearing listeners

The upper two panels of Fig. 1 show the individual results for the 12 normally hearing listeners. Results were similar for the left and right ear of each subject, so results were averaged across ears. The lower of the two panels shows results for condition I and the upper shows results for

condition II. The one error bar indicates the typical variability within a listener; it was derived by calculating the standard deviation (SD) of the estimates for each listener and then averaging the SDs across listeners. The SDs actually tended to increase with the mean value of delta E; SDs were typically about 0.18 times the value of delta E. For this reason, averages across listeners are based on geometric means.

The variability across listeners is less than found by Miller and Heise (1950). They reported interquartile ranges. The upper end of the range at each frequency was a factor of 3–4 above the lower end of the range. For example, at 1000 Hz, the frequency difference at the trill threshold had an interquartile range extending from about 100–340 Hz. In our results, the upper end of the interquartile range was always less than a factor of 1.9 larger than the lower end. On average, the upper end was a factor of 1.5 larger than the lower end. The reduced variability found by us can probably be attributed to the specific instructions given to the subject, to the fact that the fission boundary was always approached from a large frequency separation, and to the greater number of estimates contributing to the mean for each subject (Miller and Heise obtained only two judgments for each condition).

To assess the stability and reproducibility of the results, four of the listeners were retested after an interval ranging from 7–18 months. For each listener and each condition, the difference in mean delta E value for the two test periods was expressed as a proportion of the mean for the two periods. These proportions were then averaged, ignoring the sign of the difference. The resulting average proportions were 0.10, 0.15, 0.09, and 0.25 for the four subjects. Put another way,

TABLE IV. As in Table III but for the bilaterally hearing-impaired listeners.

| | Frequency (Hz) | | | | | | | |
| | 250 | | 500 | | 1000 | | 2000 | |
Ear (r=right, l=left)	r	l	r	l	r	l	r	l
RMS (35)	15/57	23/63	22/62	20/60	51/74	57/78	60/82	61/83
AVH (45)	67/86	62/83	62/81	53/76	69/85	60/80	75/91	69/87
PMA (62)	65/85	66/85	52/75	51/75	48/72	64/79	63/84	87/98
EEG (62)	38/68	44/72	45/71	45/71	48/72	47/72	47/74	47/74
EHT (65)	63/83	58/80	70/86	78/91	69/85	87/96	70/88	82/95
HAC (73)	58/80	45/73	52/75	46/69	79/91	43/69	83/96	61/83
RFI (75)	41/70	40/70	68/85	61/81	70/86	75/89	76/92	75/91
EWG (78)	56/79	57/80	60/80	54/76	72/87	73/87	77/93	71/89
JHC (78)	47/72	46/71	49/74	50/75	49/74	48/73	60/85	59/84
VDH (79)	65/85	57/80	63/82	51/75	62/81	44/70	68/87	58/81

M. M. Rose and B. C. J. Moore: Perceptual grouping
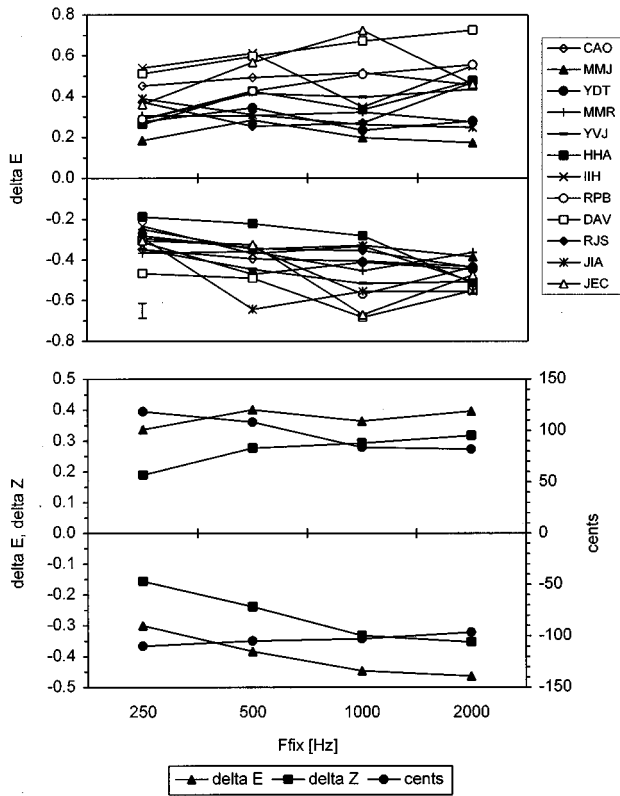
FIG. 1. The upper two panels show individual results for the twelve normally hearing listeners. The frequency difference between tones A and B at the fission boundary has been converted to delta E and plotted as a function of the frequency of tone A, F_{fix}. The lower of the two panels shows results for condition I and the upper panel shows results for condition II. The error bar indicates the average within-subject standard deviation of the delta E magnitudes. The lower two panels show geometric mean values plotted as delta E (triangles), delta Z (barks—squares), and cents (circles).

FIG. 2. Individual results for the four unilaterally hearing-impaired listeners. The left panels show the fission boundaries, expressed as delta E magnitudes, and the right panels show the absolute thresholds for each listener at the four frequencies used for tone A (F_{fix}).

the scores taken on two widely separate occasions typically differed by about 15%.

The lower two panels in Fig. 1 show the geometric mean results across listeners, plotted in three different ways. The circles show the frequency ratios of tones A and B at the fission boundary, plotted in cents (hundredths of a semitone). Miller and Heise (1950) found that the trill threshold was almost constant when expressed in this way, at least for frequencies up to 4000 Hz. The trill threshold was typically between 160 and 260 cents. Our values of the fission boundary also vary little as a function of frequency when expressed in cents, typically being about ±100 cents, although there is a slight trend for thresholds to decrease with increasing frequency; the ratio of the mean threshold at 250 Hz to the mean threshold at 2000 Hz is 1.25, averaged across conditions I and II. Our thresholds are smaller than those of Miller and Heise, probably because of our specific instructions to the listeners to try to hear fission as long as possible.

The triangles show the mean thresholds expressed in terms of delta E. Again, the thresholds vary little with frequency when expressed in this way, although there is a trend for the delta E magnitudes to increase with increasing frequency; the ratio of the mean threshold at 2000 Hz to the mean threshold at 250 Hz is 1.38, averaged across conditions I and II. The squares show the thresholds expressed in Barks

(Zwicker, 1961), using the equation for the Bark scale proposed by Traunmüller (1990). In this case, the variation of the thresholds with frequency is somewhat more distinct; the ratio of the mean threshold at 2000 Hz to the mean threshold at 250 Hz is 1.97.

The results are reasonably consistent with the prediction of the model of Beauvois and Meddis (1996) that the delta E magnitudes should be constant across frequency. Thresholds are more nearly constant when expressed in this way than when expressed in Barks. However, the thresholds are also nearly constant when expressed in cents. Further work using even lower signal frequencies would be necessary to decide whether one type of scale is clearly preferable to another; the cent, delta E and Bark values for a given frequency difference deviate more and more as the frequency is decreased below 250 Hz.

2. Unilaterally hearing-impaired listeners

Figure 2 shows the individual results for the four unilaterally hearing-impaired listeners. The right panels show the absolute thresholds for each listener at the four frequencies used for tone A; these were obtained using the forced-choice task, but converted to dB HL using the correction factors in Table II. The left panels show the fission boundaries for the impaired ear (dark bars) and the normal ear (open bars). Error bars indicate ± one standard deviation across the means of three runs.

The delta E magnitudes are not obviously larger for the impaired ears than for the normal ears, and the opposite is true for some listeners at some frequencies. This is *not* consistent with the prediction of the model of Beauvois and Meddis (1996), because one would expect broader filters in the impaired ears, and therefore larger delta E magnitudes at the fission boundary.

The inter-listener variability is rather high, with delta E magnitudes ranging from about 0.2 (listener VVF) to 1.4 (listener BOS; note the larger range of ordinate values for this listener). It is curious that listener BOS has larger delta E magnitudes than the normal listeners even in her normal ear. The average magnitude of delta E for her normal ear is 0.85 (SD 0.15). The largest average magnitude found for any individual in the group with normal hearing in both ears was 0.66 (left ear only, SD=0.12; mean over both ears: 0.58, SD=0.13). The delta E magnitudes of the other three listeners are generally in the same range as those obtained from normally hearing listeners, although the individual standard deviations are large in some conditions for listeners AFW and VVF. The large standard deviations occur mainly for the impaired ears.

For the lowest frequency, VVF could not do the task at all when the stimuli were presented to the impaired ear and the variable tone started lower in frequency than the 250-Hz fixed tone; this case is indicated in the panel as n.p. (not possible). It appears that this listener had difficulty either in hearing very low-frequency tones, or in perceiving their pitches. This may have happened because there was a region at the apical end of the cochlea with no functioning inner hair cells or neurons, a so-called ''dead'' region (Thornton and Abbas, 1980; Florentine and Houtsma, 1983; Moore and Glasberg, 1997; Moore *et al.*, 1997b). In fact, in an earlier study with this same listener it had been suggested that she had a ''dead'' region for characteristic frequencies up to 280 Hz (Moore and Glasberg, 1997).

The results of AFW may also have been influenced by a ''dead'' region. In his impaired ear, the results were unusually variable when tone B approached a fixed 250-Hz tone from above, or when tone B approached a fixed 500-Hz tone from below. It is possible that he had a ''dead'' region for characteristic frequencies in the range 250–500 Hz. This is also broadly consistent with the ''dead'' region estimated in an earlier study using this same listener (Moore and Glasberg, 1997), although in that study the dead region was estimated to extend to somewhat higher frequencies.

For two of the listeners, BOS and AFW, the ERBs of the auditory filters in each ear had been measured in a previous study using an abbreviated version of the notched-noise method (Stone *et al.*, 1992), for center frequencies of 500, 1000, and 2000 Hz. For the impaired ear of AFW at 500 Hz, the threshold for detecting the signal did not decrease with increasing notch width, consistent with the hypothesis that he had a dead region around 500 Hz. In all other cases it was possible to estimate the ERBs of the auditory filters. Figure 3 shows a scatter plot of the delta E magnitudes versus the ERBs. The delta E magnitudes were averaged for conditions I and II at each center frequency. The ERBs were expressed relative to the ERB for normal listeners at each center fre-
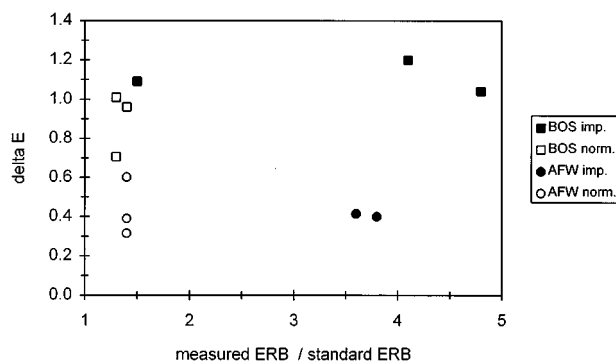
FIG. 3. Delta E magnitudes for listeners BOS and AFW, averaged for conditions I and II and plotted as a function of the ERB of the auditory filter at the corresponding center frequency. The ERB values are expressed relative to the ''normal'' ERB at each frequency.

quency, using the equation given by Glasberg and Moore (1990). Note that the noise level used to measure the ERBs was rather high, which accounts for the fact that the ERBs are slightly larger than ''normal'' even for the normal ears; the equation given by Glasberg and Moore applies for moderate noise levels. The ERBs for the impaired ears are clearly larger than those for the normal ears, except for BOS at a center frequency of 1000 Hz. However, the delta E magnitudes are not clearly larger for the impaired than for the normal ears. For both listeners, there is no clear relationship between the delta E magnitudes and the ERBs.

In summary, estimates of the fission boundary may be unusually variable, or even impossible to determine, when one of the tones falls in a ''dead'' region. Otherwise, there is no clear indication of differences in delta E magnitudes between the normal and impaired ears, even in cases where direct measurements indicate that the auditory filters are broader in the impaired ears. This is not consistent with the predictions of the model of Beauvois and Meddis (1996).

3. Bilaterally hearing-impaired listeners

Figures 4–6 show the individual results for the ten bilaterally hearing-impaired listeners. As in Fig. 2, the delta E magnitudes are shown in the left panels, and the audiograms in the right. One obvious finding is the large inter-listener variability, with delta E magnitudes at individual frequencies ranging from 0.1 (listener RFI, Fig. 4) to 2.5 (listener AVH, Fig. 6). This variability makes it necessary to plot the results on three different scales, and it makes it unwise to average across listeners.

About half of the bilaterally impaired listeners show delta E magnitudes at the fission boundary which are similar to those obtained for the normally hearing listeners, despite hearing losses of up to 70 dB HL. With this amount of hearing loss, it is unlikely that these listeners have auditory filters which are as finely tuned as the filters in normal ears (Pick *et al.*, 1977; Glasberg and Moore, 1986; Moore, 1995). Again, this is not consistent with the prediction of the model of Beauvois and Meddis (1996). However, some of the bilaterally impaired listeners do show very much larger delta E
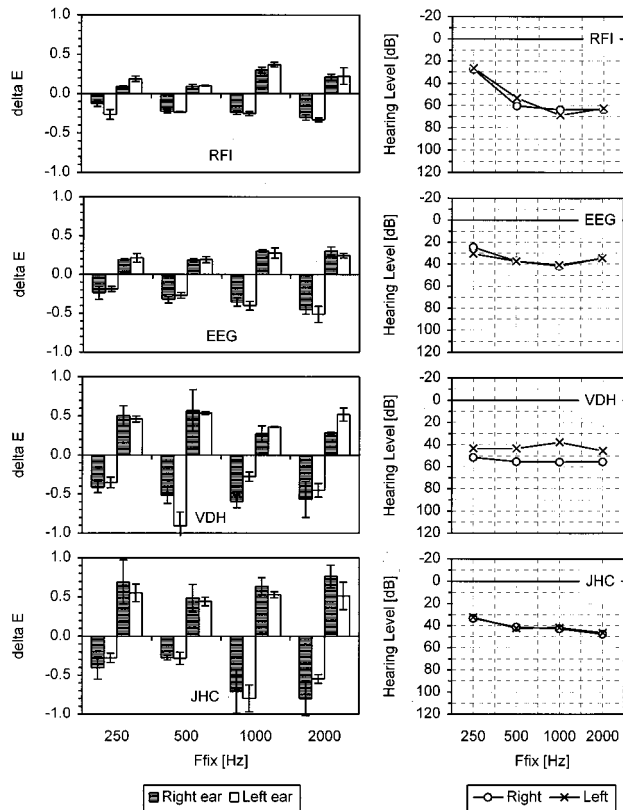
FIG. 4. As in Fig. 2, but showing results for four bilaterally hearing-impaired listeners.
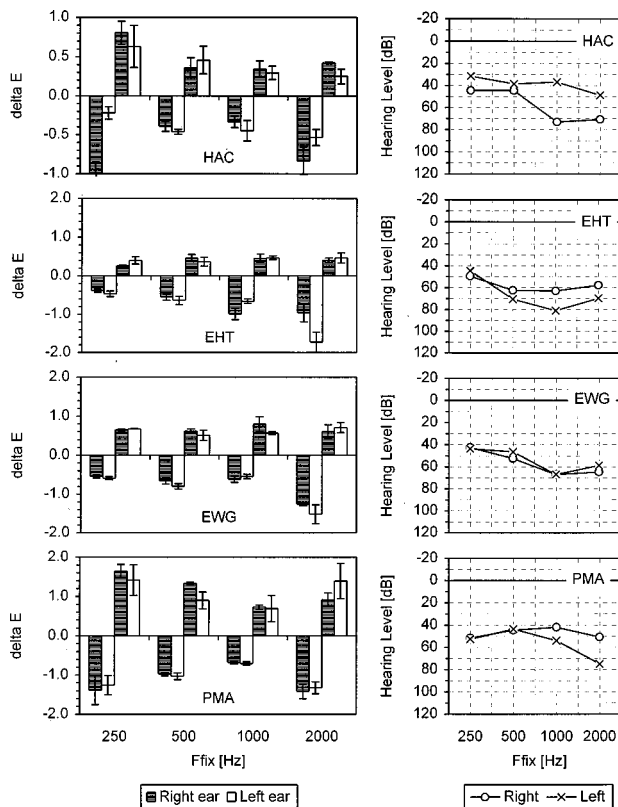
FIG. 5. As in Fig. 4. Note that the ordinate covers a greater range than in Fig. 4 for subjects EHT, EWG, and PMA.
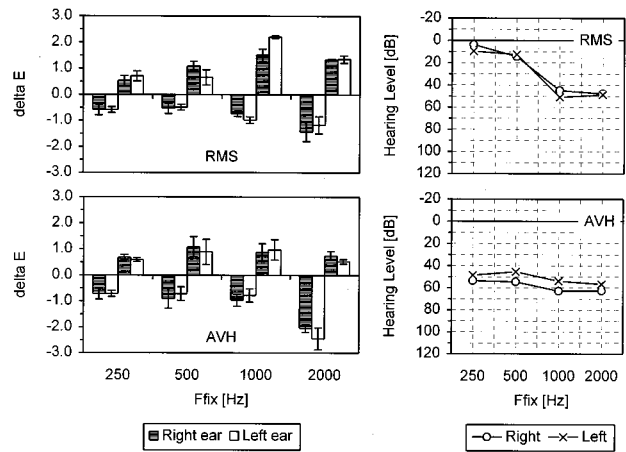
FIG. 6. As in Fig. 4 but for two subjects only. Note that the ordinate covers a greater range than in Figs. 4 and 5.

magnitudes at the fission boundary, as predicted by their model.

A further way of testing the model of Beauvois and Meddis (1996) is to see whether the pattern of delta E magnitudes across frequency is related to the pattern of hearing loss across frequency. Previous studies have shown that the width of the auditory filter is correlated with the absolute threshold, although the correlation is sometimes not very high (Pick *et al.*, 1977; Glasberg and Moore, 1986; Laroche *et al.*, 1992; Moore, 1995; Stone, 1995); higher absolute thresholds are associated with broader auditory filters. Therefore, if the model of Beauvois and Meddis is correct, higher absolute thresholds should be associated with larger delta E magnitudes. This prediction is best tested by studying the results of individual listeners, since then the decision criterion used should not influence the results.

We will consider several individual cases. The expectation of an association between absolute threshold and the delta E magnitudes is roughly borne out in two cases: RFI (Fig. 4) and RMS (Fig. 6). RFI has a mild hearing loss at low frequencies, increasing to moderate at medium and high frequencies. His delta E magnitudes are unusually small at low frequencies (compared to normally hearing listeners), and increase somewhat at 1000 and 2000 Hz. RMS has normal delta E magnitudes at 250 Hz, where her absolute threshold is nearly normal. At 1000 Hz, where her hearing loss is about 50 dB, the delta E magnitudes are between two and four times as large as normal.

For other listeners, there are clear deviations from the predictions of the model of Beauvois and Meddis. For example, listener HAC (Fig. 5) has absolute thresholds that tend to worsen with increasing frequency, but his delta E magnitudes are largest for the lowest frequency tested. Several listeners have almost flat hearing losses, but their delta E magnitudes vary markedly with frequency (e.g., listener AVH in Fig. 6 and listener EHT in Fig. 5).

To provide a more rigorous test of the predictions of the model of Beauvois and Meddis, we "normalized" the delta E magnitudes in an attempt to eliminate effects of differences in criterion across listeners. The mean value of delta E across frequency and across conditions I and II was determined for each ear of the bilaterally impaired listeners. All

delta E magnitudes for that ear were then expressed as a proportion of the mean value of delta E. These normalized magnitudes reflect the pattern of changes across frequency. Similarly, the absolute threshold values were normalized to reflect changes across frequency; the mean audiometric threshold for the frequencies 250, 500, 1000, and 2000 Hz was determined for each ear of each listener, and the threshold at each frequency was then expressed as a deviation from that mean. There was a relationship between the normalized delta E magnitudes and the normalized audiometric thresholds. The correlation was modest ($r = 0.5$), but was highly significant ($n = 80$, $p < 0.001$). This correlation provides some support for the model of Beauvois and Meddis. However, the correlation may have occurred because some factor other than the auditory filter bandwidth influences perceptual stream formation, and that factor is also correlated with the absolute threshold.

III. DISCUSSION

The results for the normally hearing listeners showed less variability across listeners than the results of Miller and Heise (1950). Also, repeated measurements on four listeners showed a high degree of stability and reproducibility. However, considerable variability across listeners remained. This can probably be attributed to differences in response criterion across listeners, and possibly also to musical training.

The mean delta E magnitudes for the normally hearing listeners were reasonably constant across frequency, which suggests that the E scale corresponds reasonably well to the internal representation of frequency and is consistent with the model of Beauvois and Meddis (1996). However, there was a slight trend for delta E magnitudes to increase with increasing frequency.

The results for the unilaterally hearing-impaired listeners showed no clear difference in the fission boundary for the normal and impaired ears for any listener. For two of these listeners, estimates of the ERBs of their auditory filters using an abbreviated notched-noise method revealed filters that were generally much broader for the impaired than for the normal ears. These results are *not* consistent with the prediction of the model of Beauvois and Meddis (1996). However, there were a few cases where the fission boundary was difficult to determine and/or highly variable in the impaired ear. This appears to be associated with dead regions in the cochlea where there are no functioning inner hair cells and/or neurons.

Half of the bilaterally hearing-impaired listeners did show abnormally large delta E magnitudes at the fission boundary. However, within individual listeners, the pattern of the delta E magnitudes was not always clearly related to the pattern of absolute thresholds across frequency. An analysis based on normalized delta E magnitudes, which allowed for the effects of individual differences in criterion, revealed a moderate association between absolute threshold and delta E magnitudes.

Overall, these results lend only modest support to the model of Beauvois and Meddis (1996). While some hearing-impaired listeners did show abnormally large delta E magnitudes, it appears that broadening of the auditory filters is not

the only factor responsible for this. It is worthwhile, therefore, to consider what other factors might contribute to the large delta E magnitudes found for some listeners.

One possibility is that the fission boundary depends on the frequency-to-place mapping within the cochlea; the fission boundary may correspond to the point where the peaks in the excitation patterns evoked by successive tones are separated by a constant distance along the basilar membrane; the exact value of this distance might vary across listeners depending on their criterion. This could explain why the delta E magnitudes for normal hearing listeners are roughly constant across frequency; the E scale corresponds reasonably well to a scale of distance along the basilar membrane in normally hearing listeners (Moore, 1986).

Cochlear hearing impairment *can* result in a change in the frequency-to-place mapping within the cochlea (Liberman, 1982; Sellick *et al.*, 1982; Liberman and Dodds, 1984; Liberman *et al.*, 1986). Typically, the characteristic frequency (CF) for a given place decreases with increasing hearing loss, especially loss of outer hair cells. Correspondingly, the position of peak excitation produced by a tone of given frequency shifts toward the base of the cochlea as hearing loss increases. When a hearing loss is relatively uniform across frequency, all CFs may shift by a similar ratio. In such a case, delta E magnitudes may still correspond to constant distances along the BM, and the fission boundary may have normal or near normal magnitudes across frequency. However, when a hearing loss varies across frequency, and especially when the amount of outer hair cell damage varies across frequency, the frequency-to-place mapping may be distorted in form. In such a case, delta E magnitudes will not correspond to constant distances along the BM, and the delta E magnitudes at the fission boundary may vary across frequency.

Another factor that may influence the fission boundary is the clarity of the pitch sensations evoked by the tones. It may be that the pitches of successive tones need to be clearly different for fission to be heard. If the pitches of the tones are unclear, this may lead to larger-than-normal delta E magnitudes at the fission boundary. It is known that the frequency discrimination of sinusoids is worse than normal in people with cochlear hearing loss (Hall and Wood, 1984; Nelson and Freyman, 1986; Moore and Peters, 1992; Moore, 1995). Also, several of our hearing-impaired listeners reported that the pitches of the tones were unclear or ''fuzzy'' when they fell in certain frequency ranges. Frequency discrimination abilities show marked variation across hearing-impaired listeners, and are not well correlated with estimates of the auditory filter bandwidth (Moore and Peters, 1992). This may be the case because frequency discrimination (and the perception of pitch) depends on the timing of neural impulses (phase locking) as well as on the distribution of activity along the basilar membrane. Variations in frequency discrimination ability across listeners could well account for some of the individual differences in the fission boundary found by us. We are currently conducting further experiments to explore this possibility.

The large delta E magnitudes found for some of the hearing-impaired listeners suggest that these listeners would

have difficulty in the perceptual segregation of sequences of sounds; they would require larger-than-normal differences in frequency between successive sounds for segregation to be possible. The significance of this for perception in everyday life, especially for speech perception in noise, is not clear. However, it seems likely that abnormally large delta E magnitudes might lead to some difficulties in the perception of musical sounds. The largest delta E magnitudes found by us correspond to a frequency ratio between successive tones of about 1.35, which is equivalent to about five semitones. A listener requiring such a large ratio at the fission boundary might have difficulty in segregating the melodic lines played by two instruments when the instruments were playing in ranges separated by less than five semitones.

The recognition of simple melodies in the presence of competing melodies by subjects with cochlear hearing loss has been directly examined in two studies. de Laat and Plomp (1985) presented a reference melody consisting of a sequence of four tones, followed by two comparison intervals, one containing the reference melody and the other containing the same four tones but with the order of the second and third tones reversed. The comparison intervals also contained two interfering melodies in which the competing tones were synchronized to those of the reference melody. The listener's task was to select the comparison interval containing the reference melody (three-interval, two-alternative forced-choice task). They found that normally hearing subjects required a separation of about 5 semitones between the reference and competing melodies in order to perform the task reliably (79% correct), while the subjects with cochlear hearing loss required a separation of 27 semitones. There were large individual differences among the hearing-impaired subjects, but they were not correlated with absolute thresholds, or with a measure of frequency selectivity at 1000 Hz. Note that this experiment involved *simultaneous* presentation of the reference and competing melodies. In this situation, reduced frequency selectivity could have substantial effects simply as a result of simultaneous masking; this could partly account for their very large effect.

Grose and Hall (1996) used a similar task, but they also investigated the effect of introducing a temporal asynchrony between the tones in the reference melody and the tones in the competing melodies. This would be expected to enhance the perceptual segregation of the reference and competing melodies. However, it did not markedly improve performance on the task, the hearing-impaired listeners continuing to perform more poorly that the normally hearing listeners. They also investigated the effect of amplitude modulating the tones in the reference and competing melodies, in phase at a 25-Hz rate. This would be expected to promote the perceptual fusion of the reference and interfering melodies. Again, however, it had little effect on performance.

It should be noted that, in both of these studies, the tones in the reference melody were either synchronous with the tones in the competing melodies, or they started after and ended before the tones in the competing melodies; the tones in the reference melody were never presented alone in the comparison intervals. Thus the task was very different to one where reference and competing sounds are interleaved; this type of task would be more comparable to our task.

In another experiment, Grose and Hall (1996) required listeners to detect a gap between two tones. When the frequency separation between the tones was small, performance was similar for the normally hearing and hearing-impaired subjects. When the separation was increased to a large value, gap thresholds increased by a factor of 5 for the normal subjects and 9 for the impaired subjects. This increase was explained by assuming that when the tones are widely separated in frequency they are heard in separate perceptual streams, and judgments of timing across steams are difficult. If this explanation is correct, then the results suggest that streaming is not determined by the overlap of excitation patterns of the two tones. If this were the case, then the increase in gap threshold with increasing frequency separation should have been smaller for the hearing-impaired than for the normally hearing subjects, whereas the opposite was observed. The gap thresholds for the hearing-impaired subjects were not correlated with the absolute thresholds or with a simplified notched-noise measure of frequency selectivity.

Grose and Hall also investigated the effect of embedding the widely separated tone pair in a sequence of alternating tones. The idea was that the upper tone would be heard as part of a high-frequency stream and the lower tone as part of a low-frequency stream. The gap could then be detected as an irregularity of the tempo *within* one stream, which might lead to improved performance. Embedding the tone pair in the sequence of alternating tones did indeed lead to reduced gap thresholds, and the improvement was larger for the hearing-impaired than for the normally hearing group. However, performance remained worse for the hearing-impaired group. Grose and Hall concluded that ''... the (presumably central) processes responsible for stream segregation appear to be functional in listeners with long-term cochlear disorders.'' They also concluded that ''perceptual organization in the cochlear loss group appeared to require a greater frequency separation between presumed auditory streams in comparison to listeners with normal hearing.'' This is consistent with the results for some, but not all, of our bilaterally hearing-impaired subjects. However, like us, Grose and Hall noted that there were large individual differences among both the normal and the hearing-impaired subjects.

IV. CONCLUSIONS

(1) The task developed for this study gives stable and reproducible results within listeners, although variability across listeners can be substantial, even when hearing is normal.

(2) The mean delta E magnitudes at the fission boundary for the normally hearing listeners were reasonably constant across frequency, which suggests that the E scale corresponds reasonably well to the internal representation of frequency and is consistent with the model of Beauvois and Meddis (1996).

(3) The results for the unilaterally impaired listeners showed no clear difference in the fission boundary for the normal and impaired ears. This is *not* consistent with the prediction of the model of Beauvois and Meddis (1996).

(4) Half of the bilaterally hearing-impaired listeners did show abnormally large delta E magnitudes at the fission boundary. An analysis based on normalized delta E magnitudes, which allowed for the effects of individual differences in criterion, revealed a moderate association between absolute threshold and delta E magnitudes.

(5) The abnormally large delta E magnitudes at the fission boundary found for some hearing-impaired listeners may be partly related to distorted frequency-to-place mapping in the cochlea, and/or to reduced clarity of the pitches of the tones.

(6) Whatever the reasons for the large magnitudes of delta E found for some listeners, it is clear that such magnitudes *do* occur. Listeners showing large delta E magnitudes would have difficulty in forming separate perceptual streams from successive sounds even when those sounds differed markedly in frequency. This may contribute to the difficulty that they have in everyday life when multiple sound sources are active simultaneously.

ACKNOWLEDGMENTS

Beauvois, M. W., and Meddis, R. (**1996**). ''Computer simulation of auditory stream segregation in alternating-tone sequences,'' J. Acoust. Soc. Am. **99**, 2270–2280.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (Bradford Books, MIT, Cambridge, MA).

Darwin, C. J., and Carlyon, R. P. (**1995**). ''Auditory grouping,'' in *Hearing*, edited by B. C. J. Moore (Academic, San Diego).

de Laat, J. A. P. M., and Plomp, R. (**1985**). ''The effect of competing melodies on melody recognition by hearing-impaired and normal-hearing listeners,'' J. Acoust. Soc. Am. **78**, 1574–1577.

Fletcher, H. (**1940**). ''Auditory patterns,'' Rev. Mod. Phys. **12**, 47–65.

Florentine, M., and Houtsma, A. J. M. (**1983**). ''Tuning curves and pitch matches in a listener with a unilateral, low-frequency hearing loss,'' J. Acoust. Soc. Am. **73**, 961–965.

Glasberg, B. R., and Moore, B. C. J. (**1986**). ''Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments,'' J. Acoust. Soc. Am. **79**, 1020–1033.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hearing Res. **47**, 103–138.

Greenwood, D. D. (**1961**). ''Critical bandwidth and the frequency coordinates of the basilar membrane,'' J. Acoust. Soc. Am. **33**, 1344–1356.

Greenwood, D. D. (**1990**). ''A cochlear frequency-position function for several species—29 years later,'' J. Acoust. Soc. Am. **87**, 2592–2605.

Grose, J. H., and Hall, J. W. (**1996**). ''Perceptual organization of sequential stimuli in listeners with cochlear hearing loss,'' J. Speech Hear. Res. **39**, 1149–1158.

Hall, J. W., and Wood, E. J. (**1984**). ''Stimulus duration and frequency discrimination for normal-hearing and hearing-impaired subjects,'' J. Speech Hear. Res. **27**, 252–256.

ISO 389-7 (**1996**). *Acoustics-Reference zero for the calibration of audiometric equipment. Part 7: Reference threshold of hearing under free-field and diffuse-field listening conditions* (International Organization for Standardization, Geneva).

Laroche, C., Hétu, R., Quoc, H. T., Josserand, B., and Glasberg, B. (**1992**). ''Frequency selectivity in workers with noise-induced hearing loss,'' Hearing Res. **64**, 61–72.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Liberman, M. C. (**1982**). ''The cochlear frequency map for the cat: Labeling auditory-nerve fibers of known characteristic frequency,'' J. Acoust. Soc. Am. **72**, 1441–1449.

Liberman, M. C., and Dodds, L. W. (**1984**). ''Single neuron labeling and chronic cochlea pathology. III. Stereocilia damage and alterations in threshold tuning curves,'' Hearing Res. **16**, 54–74.

Liberman, M. C., Dodds, L. W., and Learson, D. A. (**1986**). ''Structure-function correlation in noise-damaged ears: A light and electron-microscopic study,'' in *Basic and Applied Aspects of Noise-Induced Hearing Loss*, edited by R. J. Salvi, D. Henderson, R. P. Hamernik, and V. Colletti (Plenum, New York).

Miller, G. A., and Heise, G. A. (**1950**). ''The trill threshold,'' J. Acoust. Soc. Am. **22**, 637–638.

Moore, B. C. J. (**1986**). ''Parallels between frequency selectivity measured psychophysically and in cochlear mechanics,'' Scand. Audiol. **25**, 139–152.

Moore, B. C. J. (**1995**). *Perceptual Consequences of Cochlear Damage* (Oxford U.P., Oxford).

Moore, B. C. J. (**1997**). *An Introduction to the Psychology of Hearing* (Academic, San Diego), 4th ed.

Moore, B. C. J., and Glasberg, B. R. (**1983**). ''Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,'' J. Acoust. Soc. Am. **74**, 750–753.

Moore, B. C. J., and Glasberg, B. R. (**1997**). ''A model of loudness perception applied to cochlear hearing loss,'' Aud. Neurosci. **3**, 289–311.

Moore, B. C. J., and Peters, R. W. (**1992**). ''Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity,'' J. Acoust. Soc. Am. **91**, 2881–2893.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997a**). ''A model for prediction of thresholds, loudness and partial loudness,'' J. Audio Eng. Soc. **45**, 224–240.

Moore, B. C. J., Wojtczak, M., and Vickers, D. A. (**1996**). ''Effect of loudness recruitment on the perception of amplitude modulation,'' J. Acoust. Soc. Am. **100**, 481–489.

Moore, B. C. J., Vickers, D. A., Glasberg, B. R., and Baer, T. (**1997b**). ''Comparison of real and simulated hearing impairment in subjects with unilateral and bilateral cochlear hearing loss,'' Br. J. Audiol. (in press).

Nelson, D. A., and Freyman, R. L. (**1986**). ''Psychometric functions for frequency discrimination from listeners with sensorineural hearing loss,'' J. Acoust. Soc. Am. **79**, 799–805.

Pick, G., Evans, E. F., and Wilson, J. P. (**1977**). ''Frequency resolution in patients with hearing loss of cochlear origin,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London).

Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (**1982**). ''Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique,'' J. Acoust. Soc. Am. **72**, 131–141.

Shaw, E. A. G. (**1974**). ''Transformation of sound pressure level from the free field to the eardrum in the horizontal plane,'' J. Acoust. Soc. Am. **56**, 1848–1861.

Stone, M. A. (**1995**). ''Spectral enhancement for the hearing impaired,'' Ph.D. thesis, University of Cambridge.

Stone, M. A., Glasberg, B. R., and Moore, B. C. J. (**1992**). ''Simplified measurement of impaired auditory filter shapes using the notched-noise method,'' Br. J. Audiol. **26**, 329–334.

Thornton, A. R., and Abbas, P. J. (**1980**). ''Low-frequency hearing loss: Perception of filtered speech, psychophysical tuning curves, and masking,'' J. Acoust. Soc. Am. **67**, 638–643.

Traunmüller, H. (**1990**). ''Analytical expressions for the tonotopic sensory scale,'' J. Acoust. Soc. Am. **88**, 97–100.

van Noorden, L. P. A. S. (**1971**). ''Rhythmic fission as a function of tone rate,'' IPO Ann. Prog. Rep. **6**, 9–12.

van Noorden, L. P. A. S. (**1975**). ''Temporal coherence in the perception of tone sequences,'' Ph.D. thesis, Eindhoven University of Technology.

Zwicker, E. (**1961**). ''Subdivision of the audible frequency range into critical bands (Frequenzgruppen),'' J. Acoust. Soc. Am. **33**, 248.

Increment and decrement detection in sinusoids as a measure of temporal resolution

Andrew J. Oxenham[a]
Institute for Perception Research (IPO), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Measuring thresholds for the detection of brief decrements in the level of a sinusoid is an established method of estimating auditory temporal resolution. Generally, a background noise is added to the stimulus to avoid the detection of the ''spectral splatter'' introduced by the decrement. Results are often described in terms of a temporal-window model, comprising a band-pass filter, a compressive nonlinearity, a sliding temporal integrator, and a decision device. In this study, thresholds for increments, as well as decrements, in the level of a 55 dB SPL, 4-kHz sinusoidal pedestal were measured as function of increment and decrement duration in the presence of a broadband background noise ranging in spectrum level from -20 to $+20$ dB SPL. Thresholds were also measured using a 55-dB, 8-kHz pedestal in the absence of background noise. Thresholds for decrements, in terms of the dB change in level (ΔL), were found to be more dependent on duration than those for increments. Also, performance was found to be dependent on background-noise level over most levels tested. Neither finding is consistent with the predictions of the temporal-window model or other similar models of temporal resolution. The difference between increment and decrement detection was more successfully simulated by using a decision criterion based on the maximum *slope* of the temporal-window output. © *1997 Acoustical Society of America.* [S0001-4966(97)05909-2]

PACS numbers: 43.66.Mk, 43.66.Dc, 43.66.Ba [JWH]

INTRODUCTION

Temporal resolution in the auditory system is often described in terms of the ability to detect brief fluctuations in the level of a stimulus. Many models of temporal resolution assume that the ability to detect such fluctuations is limited by a smoothing process in the auditory system. This process has often been incorporated within a more complete model of temporal resolution comprising a band-pass filter, a rectifier and power-law nonlinearity, a smoothing device (implemented as a low-pass filter or a sliding temporal integrator), and a decision device (Buunen and van Valkenburg, 1979; Viemeister, 1979; Buus and Florentine, 1985; Forrest and Green, 1987; Moore *et al.*, 1988; Plack and Moore, 1991).

Recently, measuring thresholds for brief decrements in the level of a sinusoidal pedestal has become a popular method of estimating temporal resolution, as a function of pedestal frequency (Moore *et al.*, 1993b) and level (Peters *et al.*, 1995; Moore *et al.*, 1996). In these studies, the results from the experiments have been used to derive the parameters of the smoothing device (implemented as a sliding temporal integrator or temporal window) and the subsequent decision device. It is assumed that detection of a decrement occurs when the decrease in the output level of the temporal window equals or exceeds a certain criterion level (in dB), referred to as ΔO. A stated assumption of this model is that increments are detected in essentially the same manner as decrements, namely when the increase in output of the temporal window exceeds the same criterion level (ΔO). However, this assumption has not been directly tested. In all pre-

vious temporal-window studies of increment and decrement detection (Moore *et al.*, 1993b, 1996; Peters *et al.*, 1995) thresholds for only one increment duration were measured, so it is not clear whether the same model parameters that describe decrement detection can also account for increment detection as a function of duration. Conversely, it is not clear whether an experiment measuring increment detection as a function of duration would produce the same model parameters as a decrement-detection experiment.

Studies of increment and decrement detection in broadband noise have found that very brief decrements, when measured in terms of the dB change in level (ΔL), are generally less detectable than brief increments of the same duration (Irwin and Purdy, 1982; Forrest and Green, 1987). Irwin and Purdy (1982) suggested that the data could be explained if neural adaptation was taken into account (Smith, 1979), although this approach seems not to have been pursued since. One previous study has investigated the detection of increments and decrements in the level of a 1-kHz sinusoid (de Boer, 1986). There it was also found that decrements were less detectable than increments at the shortest durations. No quantitative modeling was attempted, however. Also, the increments and decrements used by de Boer (1986) were gated on and off instantaneously. While the pedestal was presented in a background notched noise, the large bandwidth of the notch (2 octaves) makes it difficult to rule out cues associated with ''spectral splatter'' due to the onset and offset of the signal (e.g., Leshowitz and Wightman, 1971).

Macmillan (1971) provided data which indicate that the detection of brief increments and decrements may be mediated by a ''change detector,'' rather than an energy detector.

[a]Electronic mail: oxenham@ipo.tue.nl

He reached this conclusion based on the fact that the discriminability of increments and decrements is less than would be predicted based on their respective detectability. While some of his results may have been due to spectral artifacts (Leshowitz and Wightman, 1972), a further study, controlling for spectral effects, confirmed the trend of the initial findings (Macmillan, 1973). Such a change detector is inconsistent with the assumptions of the temporal-window model, and other models of temporal resolution (e.g., Buus and Florentine, 1985), which assume that detection is due to a maximum in the peak or dip of a smoothed representation of stimuli. However, another decision device, proposed by Viemeister (1979) within a model accounting for amplitude modulation detection, may be compatible with the notion of a change detector: in this model, the standard deviation of the output, regardless of the direction of the change, is used as the decision device.

One purpose of this study is to compare the detection of increments and decrements directly by measuring thresholds for both as a function of duration in the same listeners. The results are used to evaluate the ability of current models of temporal resolution to account for the data. A high pedestal frequency (4 kHz) was chosen in order to allow the measurement of relatively short increments and decrements without the bandwidth of the signal exceeding the bandwidth of the auditory filter at 4 kHz.

A second question studied here is the effect of adding a broadband background noise to the pedestal. This has been done in most previous studies in order to reduce the possibility that listeners detect the spectral splatter associated with the onset and offset of the signal. It is generally assumed that the level of the background noise can be set sufficiently high to mask any off-frequency signal energy, but not so high as to influence on-frequency performance. Purely in terms of the overall noise energy within and beyond the auditory filter centered at the signal frequency, this assumption is reasonable. However, other aspects of the background noise, such as its inherent fluctuations, may play a significant role, either within or across different frequency channels. The effect of adding noise is investigated here by measuring performance over a wide range of background-noise levels. If the assumptions of the temporal-window model are correct, performance should remain constant over a considerable portion of the range of noise levels tested.

I. EXPERIMENT 1. INCREMENT AND DECREMENT DETECTION IN BROADBAND BACKGROUND NOISE

A. Stimuli

Thresholds for detecting increments and decrements in the level of a 4-kHz sinusoidal pedestal were measured as a function of signal (increment or decrement) duration. All stimuli were generated and controlled digitally on a Silicon Graphics Indigo workstation and were played out via the built-in D/A converter at a sampling rate of 32 kHz. The pedestal had a total duration of 500 ms, gated with 50-ms raised-cosine ramps, and was presented three times in each trial, separated by 200-ms interstimulus intervals. The level

of the pedestal was always 55 dB SPL. The signal was added to one of the three pedestals at random and was temporally centered in the pedestal. The signal was either added in phase with the pedestal, to produce an increment in the overall level, or in antiphase, to produce a decrement. The signal had a steady-state duration of 2, 6, 14, and 198 ms, and was gated with 2-ms raised-cosine ramps, producing increments and decrements with half-amplitude durations of 4, 8, 16, and 200 ms.

The broadband (0–15 kHz) Gaussian background noise was generated at the beginning of each run and was stored in a 4-s circular buffer. Three 700-ms segments of the noise were chosen at random (with replacement) and added to each of the three pedestals, with each noise onset 200 ms before each pedestal onset. The three segments were then concatenated to produce a single trial of 2.1-s duration. As the noise was broadband and was gated abruptly, there was no discontinuity between the 700-ms segments and so the noise provided a continuous (2.1-s) background during each trial. Thresholds for the increments and decrements were measured for noise spectrum levels of −20, −10, 0, 10, and 20 dB SPL and were also measured in the absence of noise.

The stimuli from the Silicon Graphics D/A converter were passed through a Tucker-Davies programmable attenuator (TDT PA4) and headphone buffer (TDT HB6) before being presented to one ear of the listener via a Beyer DT 990 headset.

B. Procedure

Thresholds were measured using a three-alternative forced-choice method with a three-down one-up adaptive procedure that tracks the 79.4% correct point on the psychometric function. Each listener was tested in 2-h sessions, divided into four blocks of runs of approximately 20 min each. Within each block only increments or only decrements were tested.

At the beginning of each run the level of the signal was set to the same level as the pedestal. This resulted in a complete gap in the case of decrements and a 6-dB increase in level in the case of increments. The task of the listener was to select which of the three intervals contained the increment or decrement. Responses were made via a computer keyboard, and feedback was provided via a computer monitor. In each run the signal level was decreased after three consecutive correct responses and was increased after each incorrect response. A change from increasing to decreasing signal level, or vice versa, defines a reversal. Initially the signal step size was 8 dB. This value was halved after every second reversal until it reached its minimum value of 2 dB, after which it remained constant. A run was terminated after 14 reversals and the threshold was defined as the median signal level at the last 10 reversals. For the decrements, the signal level was not permitted to exceed 0 dB, relative to the level of the pedestal. If the adaptive procedure required a higher level, the level was maintained at 0 dB and testing continued. If 30 trials within a run required a nominal level greater than 0 dB, the run was terminated. Any runs which included nominal levels of more than 0 dB within the test phase (last 10 reversals) were discarded. This procedure was

designed to allow listeners some practice with conditions in which they initially could not detect the decrement.

Three threshold estimates were obtained for each listener and condition, and transformations of the mean of the three estimates are reported here. In the case of decrements, if more than one of the three estimates was discarded (see above), that condition was deemed not to be detectable. If only one of the three estimates had been discarded, a further estimate was made. If this was also discarded, the condition was deemed not to be detectable. If not, the mean of the three remaining conditions was recorded. In practice, conditions where the decrement was initially inaudible generally remained so. Thus a further estimate was necessary for only one data point of one listener.

Listeners were tested individually in a single-walled sound-attenuating booth, situated in a sound-attenuating room.

C. Subjects

Four normally hearing listeners participated as subjects. One was the author. The other three were male students who were paid an hourly wage for their participation. The ages of the listeners ranged from 22 to 27 years. None of the three students had previous experience in psychoacoustic tasks, and all four listeners were given 2-h practice, divided equally between increments and decrements, before data were collected. All listeners had audiometric thresholds of 15 dB HL or less at octave frequencies between 250 and 8000 Hz in the test ear. No practice effects were noted during the course of the experiment.

D. Results

Results were reasonably consistent both within and across listeners. Typically, standard deviations across the three threshold estimates for each condition and listener were less than 2 dB. Figure 1 shows the individual results from the four listeners. For each data point, the mean signal level at threshold, together with ± 1 standard error of the mean, was calculated and then transformed into the level of the increment or decrement, ΔL, defined as $20 \log[(p+\Delta p)/p]$, where p is the sound pressure of the pedestal, and Δp and $-\Delta p$ are the sound pressures of the signal for increments and decrements, respectively. As performance was fairly similar across listeners, mean thresholds across listeners were also calculated and then transformed, as described above. These mean data are shown in Fig. 2.

Consider first the effect of background noise in Fig. 2. For the decrements, thresholds seem to be dependent on background-noise level across the entire range of levels tested, especially at the shorter durations. For the increments, the effect seems less pronounced at the lower noise levels. However, increment thresholds for noise levels above -10-dB spectrum level are also strongly dependent on background-noise level. A comparison of increment with decrement thresholds in units of ΔL shows that at the longest signal duration (200 ms), thresholds seem fairly symmetric. This is consistent with the assumption of the temporal-window model that the criterion level, ΔO, is the same
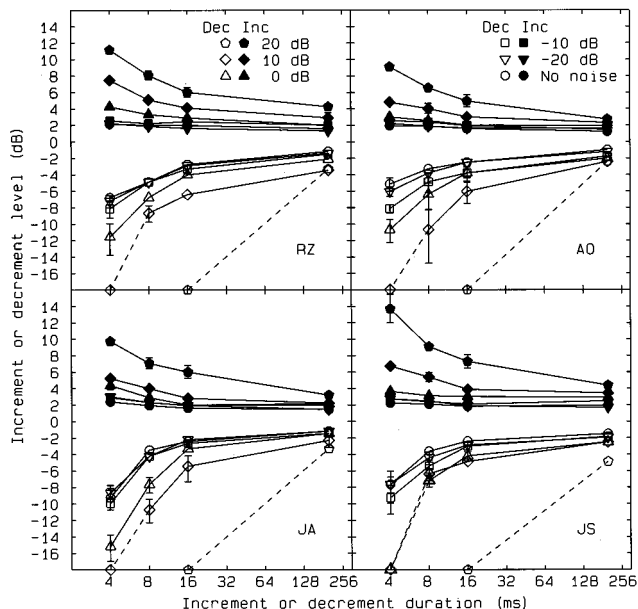
FIG. 1. Individual thresholds for detecting increments (filled symbols) and decrements (open symbols) in the level of a 55-dB, 4-kHz pedestal, in terms of ΔL. Error bars represent ± 1 standard error of the mean and are omitted if smaller than the symbol. Different symbols represent different background-noise spectrum levels, as shown in the upper panels. Symbols lying on the abscissa and joined by dashed lines represent conditions where no threshold could be measured.

whether the task is to detect an increment or a decrement. At shorter durations, performance in these units is asymmetric, both at low and high noise levels. For instance, even with no noise, the mean increment threshold at 4 ms is 2.2 dB, while the mean decrement threshold is -7.0 dB. At the highest noise level tested (20 dB), only the 200-ms decrement was detectable at all.

The units of ΔL were chosen to be consistent with previous temporal-window studies (e.g., Moore *et al.*, 1993b;
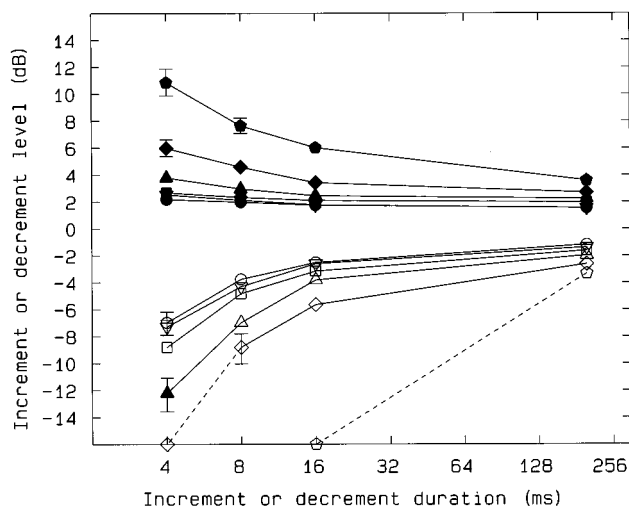
FIG. 2. Mean thresholds for detecting increments and decrements. Different symbols represent different background-noise spectrum levels, as shown in Fig. 1. Error bars represent ± 1 standard error of the mean across listeners. The single filled symbol in the decrement condition represents the mean of three, rather than four, listeners (a threshold for JS could not be measured).
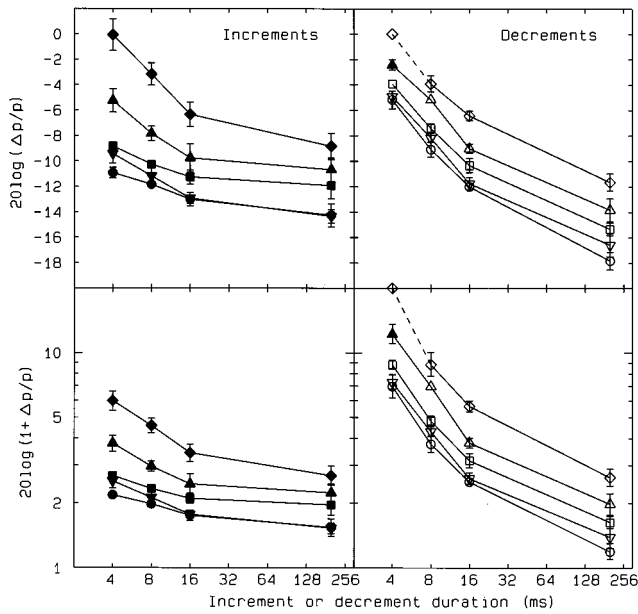
FIG. 3. Mean data from Fig. 2, replotted using two alternative level measures. The upper two panels represent increment and decrement thresholds in terms of signal level, relative to pedestal level [$20 \log(\Delta p/p)$], and the lower two panels plot ΔL, or $20 \log(1+\Delta p/p)$, on a logarithmic scale. Error bars represent ± 1 standard error of the mean.

Peters *et al.*, 1995). As pointed out by, for instance, Forrest and Green (1987), these units are highly compressive at small values. Thus small but significant differences may not be readily visible. Also, a different choice of units can affect any apparent asymmetries in the data (Forrest and Green, 1987). For this reason, the mean data from Fig. 2 are replotted in Fig. 3, using two alternative signal measures. Data from the highest noise level are omitted in this figure, as only one of the decrement conditions was measurable. The upper panels of Fig. 3 show the results in terms of signal level relative to pedestal level; the lower panels show the results in units of ΔL, but plotted on a logarithmic scale, as suggested by Buus and Florentine (1991) for level discrimination.

The asymmetry between increments and decrements at short signal durations remains apparent in Fig. 3, especially at the lower noise levels. Interestingly, in these plots the effect of the background noise is apparent for both increments and decrements, even at the longest signal duration. Another aspect, not readily apparent in Fig. 2, is that there is a tendency for thresholds for the longest decrements to be *lower* than for the longest increments, thus reversing the asymmetry at shorter durations. A similar trend was noted by Forrest and Green (1987) using broadband noise, but not by Irwin and Purdy (1982). De Boer (1986), in his study using sinusoidal pedestals, plotted his results using the same linear scale of ΔL used in Fig. 2, making comparisons between increments and decrements at long durations difficult. Nevertheless, no asymmetry at long durations is readily apparent in his data. As the plots of Fig. 3 seem to yield more information, all further figures are shown in units of ΔL on a logarithmic scale, as in the lower panels of Fig. 3.

In summary, using all three types of units, brief decrements are less detectable than brief increments of the same

duration. This conclusion is also consistent with the studies of Irwin and Purdy (1982) and Forrest and Green (1987), both of which used broadband noise as a pedestal and signal. The fact that increment thresholds at low noise levels are much less dependent on duration is also consistent with the results of Oxenham (1997). Presumably, at the highest noise levels, the noise becomes the dominant masker and temporal integration is observed which is similar to that for a signal in broadband noise alone (e.g., Gerken *et al.*, 1990). This was confirmed for listener AO: detection threshold for the 4-ms signal in the 20-dB noise alone was 59.5 dB SPL, which matches closely with the threshold signal level of 60.3 dB in the increment-detection condition with the 20-dB background noise. The pedestal itself, however, was clearly audible in all conditions.

E. Discussion

The large effect of background noise cannot be accounted for by a single-channel model, such as the temporal-window model, in which only the overall noise power within the channel is taken into account. For instance, the 10-dB noise, which renders the shortest decrement undetectable for all listeners, has an effective level of only 36 dB SPL within the equivalent rectangular bandwidth (ERB) of an auditory filter centered at 4 kHz (Glasberg and Moore, 1990). This is nearly 20 dB lower than the pedestal level (55 dB SPL), and so is not predicted to affect thresholds significantly.

One possible explanation for the effect of the background noise is that performance was based on the detection of spectral splatter at all but the highest noise levels. In this case, performance would be expected to deteriorate with increasing noise level, as the splatter would be increasingly masked.

The detectability of spectral splatter in the absence of noise can be assessed as follows. If the pedestal can be represented as a line spectrum, the phase of the signal, relative to the pedestal only affects the amplitude spectrum at the pedestal frequency (Leshowitz and Wightman, 1971). Thus all off-frequency energy due to the signal onset and offset is independent of whether the signal phase is 0 or 180 degrees. This means that the off-frequency energy for a given signal level is the same for both increments and decrements. This in turn leads to the conclusion that, even with no noise present, increment detection was *not* achieved through detection of spectral splatter. If this were the case, then increment and decrement thresholds for the shortest duration should be equal in terms of signal level (Fig. 3, upper panels). As they are not equal, it is possible to conclude that *increment* detection is based on cues other than spectral splatter.

The shortest signal employed in this experiment had a total duration of 6 ms and a 3-dB bandwidth of about 210 Hz. This is much less than the estimated ERB of the auditory filter centered at 4 kHz (456 Hz). The first side lobes of the signal's power spectrum are 17.5 dB lower in level than the main lobe and are spaced less than 1 ERB from the main lobe, at 3660 and 4340 Hz. No side lobes beyond the region of 3–5 kHz are at a level greater than -50 dB relative to the main lobe. It might therefore be expected that spectral splatter for *decrement* detection was also not detectable in the

absence of noise. Informal simulations using a gammatone filterbank (Patterson *et al.*, 1992) supported this conjecture: it was found that for filters where the output would have been above absolute threshold, the transient increase in response due to the onset and offset of the decrement was generally smaller than the minimum detectable increment in quiet. The question of the detectability of splatter is examined more empirically in experiment 3.

Whether or not thresholds for decrement detection in the absence of noise are mediated by spectral splatter, the effect of the background noise at higher levels remains to be explained. One possibility relates to the upward spread of excitation of the pedestal and signal (Zwicker, 1956). As the noise level increases, more of the pedestal and signal's excitation pattern would fall beneath masked threshold, resulting in a deterioration in performance. Moore *et al.* (1996) tested this directly in the context of increment and decrement detection. They found that performance at 4 kHz in the presence of a low-pass background noise with a cutoff frequency of 5 kHz was superior than performance in the presence of a broadband noise with the same spectrum level. However, the lowest pedestal level tested by Moore *et al.* (1996) was 70 dB SPL. This is somewhat higher than the 55 dB used here, and the benefit of the upward spread of excitation would be expected to diminish with decreasing pedestal level. Another possibility is that the inherent fluctuations of the noise directly mask, or distract from, the fluctuations introduced by the signal. In other words, the modulation frequencies introduced by the signal are masked by the modulation frequency spectrum of the noise. Much evidence has been presented recently for the role of masking in the modulation domain, both within channel (Bacon and Grantham, 1989; Houtgast, 1989; Dau *et al.*, 1997a,b; Oxenham, 1997) and across channels (Yost and Sheft, 1989; Mendoza *et al.*, 1995). The effect of the background noise was studied further in experiment 2 by measuring thresholds in noise with various spectral characteristics.

II. EXPERIMENT 2. EFFECTS OF BACKGROUND-NOISE SPECTRUM

This experiment investigated the effect of changing the spectral composition of the background noise. If the effect of the noise can be attributed to the masking of the pedestal's spread of excitation, then using a low-pass noise with a cutoff frequency near the pedestal frequency should improve performance relative to the broadband condition. Using a high-pass noise should then produce a smaller (if any) improvement. On the other hand, if on-frequency masker energy (or fluctuations) are responsible for the effect of the background noise, then using a noise with a spectral notch centered around the pedestal frequency should improve performance. Performance in these three conditions (low-pass, high-pass, and notched noise) were measured for both increment and decrement detection.

A. Method

Thresholds for increments and decrements in the 55-dB, 4-kHz pedestal were measured for half-amplitude signal durations of 4, 8, and 16 ms. The spectrum level of the back-
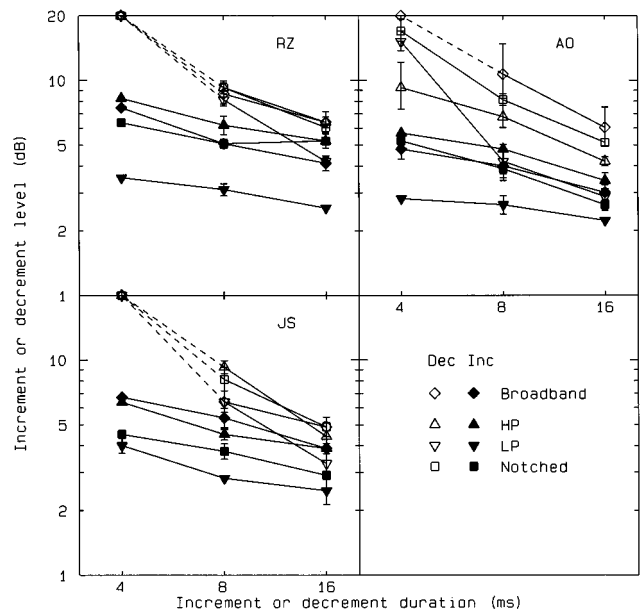
FIG. 4. Individual thresholds, in terms of ΔL, for increment and decrement detection in the presence of high-pass (upward-pointing triangles), low-pass (downward-pointing triangles), and notched noise (squares) at a spectrum level of 10 dB. Thresholds in broadband noise (diamonds) are replotted from Fig. 1. Error bars represent ± 1 standard error of the mean.

ground noise was held constant at 10 dB SPL and three different conditions were tested: low-pass noise with a cutoff frequency of 4400 Hz; high-pass noise with a cutoff frequency of 3600 Hz; and notched noise with cutoff frequencies of 3600 and 4400 Hz. The noises were generated by taking the Fourier transform of a 4-s broadband Gaussian noise buffer and setting the amplitudes of the spectral components outside the desired passband to zero. As the noise was bandlimited, concatenating two independent noise samples from the buffer, as was done in experiment 1, would have led to audible clicks at the interval boundaries. To avoid this, the noise was gated with 0.5-ms raised-cosine ramps. These were long enough to make the splatter inaudible, but not so long as to introduce audible gaps between the intervals. Thus as in experiment 1, the percept was of a 2.1-s noise present throughout a single trial. All other aspects of the stimuli and procedure were the same as in experiment 1.

Three of the four listeners from experiment 1 participated in experiment 2; listener JA was not available.

B. Results

Results from the three listeners are shown in Fig. 4. Data from the broadband condition (diamonds) are taken from Fig. 1. If the effect of the noise were due to the masking of the upward spread of excitation, performance should be better in the low-pass noise condition than in the other noise conditions. This can be seen clearly in the increment-detection data from all three listeners: the down-pointing triangles tend to be lowest. The data from the decrement-detection conditions are less clear cut. Relative to thresholds in the broadband noise, listeners RZ and JS show an im-
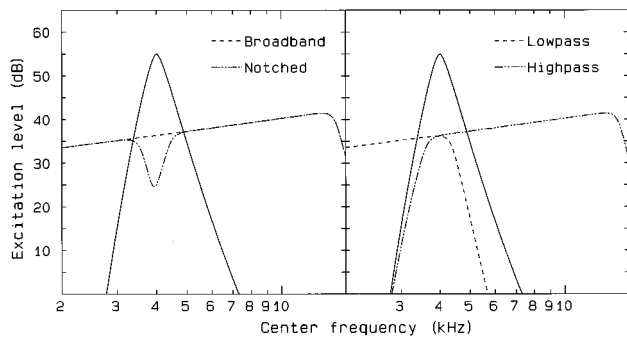
FIG. 5. Excitation patterns for the pedestal (solid curves) and the different background noises used in experiments 1 and 2.

provement in the presence of the low-pass noise only at the longest decrement duration (16 ms), while AO seems to show some improvement at all three durations. All thresholds in the presence of the low-pass noise are, however, still higher than those found with no noise present. It is in principle possible that the low-pass noise still produces some upward spread of masking. This possibility was evaluated using the excitation pattern model, proposed by Glasberg and Moore (1990). Figure 5 shows the excitation patterns of the different maskers and the pedestal. Assuming that this is a reasonable representation of the individual maskers, it seems that the upward spread of excitation due to the low-pass noise should not have a large effect on thresholds. Thus the residual effect of the low-pass masker seems not to be due to its masking of the upward spread of excitation.

The difference between thresholds in broadband noise and in the other two conditions (high-pass noise and notched noise) is dependent on the listener. For AO, decrement detection seems somewhat improved by the two new conditions, while increment detection is relatively unaffected. Conversely, listener JS shows a possible improvement due to the notched noise in the increment conditions, but shows no clear improvement in the decrement conditions. For listener RZ, performance in neither increment nor decrement conditions is improved by the notched or the high-pass noise, relative to the broadband condition.

According to the excitation pattern model, the excitation produced by the notched noise at the pedestal frequency is about 10 dB lower than that produced by the broadband noise (see Fig. 5). Therefore, if performance were limited by on-frequency noise energy, performance in the presence of notched noise should resemble that found for the 0-dB spectrum level broadband masker (Fig. 1). Instead, thresholds are generally higher, indicating that the effects of the noise are probably not solely due to on-frequency energy either.

The fact that, for listeners RZ and JS, none of the conditions substantially improved decrement detection suggests that the noise is playing a masking role which cannot be accounted for purely in terms of its spectral characteristics. It may be that the noise fluctuations in off-frequency channels mask the fluctuations introduced by the decrement, in a manner similar to that observed in modulation detection interference (MDI) (e.g., Yost and Sheft, 1989). Whatever the precise mechanisms, it is clear that a single-channel model of

temporal resolution, such as the temporal-window model, cannot account for the results.

In summary, the effects of the noise are not clearly understood and cannot be predicted by current models of auditory processing. Therefore, including a background noise when estimating temporal resolution poses the problem of which the "correct" level to use should be. It would clearly be preferable to dispense with the noise altogether. As mentioned earlier, it is likely that at 4 kHz spectral splatter was not audible, even in the absence of noise. The following experiment tests this conjecture.

III. EXPERIMENT 3. EFFECTS OF SIGNAL PHASE

In this experiment, the detectability of spectral splatter in the absence of noise was tested. As stated earlier, the detection of spectral splatter should be independent of the phase relationship between the pedestal and the signal. This is because the pedestal can be represented by a line spectrum, meaning that the pedestal and signal only interact at their common center frequency. The assumption that the detectability of splatter is independent of signal phase is tested empirically in the Appendix and is found to hold for the three listeners tested there.

The present experiment compared performance with the signal added in antiphase (180 degrees; as for the decrements in experiment 1) with performance when the signal was added with a phase shift of 120 degrees. When the signal and pedestal are at the same level, adding an antiphasic signal results in a complete gap, while adding a signal at 120 degrees results in no change of level, except during the onset and offset portions of the signal. At lower signal levels, the difference between the two conditions diminishes, but the antiphasic signal continues to produce the greater decrement. Nevertheless, both signals produce the same amount of off-frequency energy (see above and the Appendix).[1] Thus if detection is mediated by the largest temporal "dip" in an on-frequency filter, performance should be worse for the 120-degree condition than for the antiphasic (180-degree) condition. If detection is mediated by spectral splatter, both conditions should produce the same threshold signal level. Only the shortest signal duration was tested, as spectral splatter is most likely to play a role at that duration.

A. Method

Thresholds for a 4-kHz, 4-ms (half-amplitude duration) signal, temporally centered in a 4-kHz, 500-ms pedestal, were measured in the absence of background noise. The phase of the signal was either 120 or 180 degrees, relative to the pedestal. The 180-degree condition was a replication of the no-noise decrement-detection task of experiment 1. For both phases, the signal level was not permitted to exceed 0 dB, relative to the pedestal level of 55 dB SPL. All four listeners from experiment 1 participated. After the initial results were collected, the experiment was repeated using three new normally hearing listeners (for reasons discussed below) with ages of 26, 30, and 31 years.

TABLE I. Signal thresholds (dB), relative to the pedestal level, for the detection of a signal with a phase of either 120 or 180 degrees, relative to the pedestal. Estimated standard deviations are given in parentheses.

Listener	180 degrees	120 degrees
AO	−7.3 (1.2)	−3.0 (1.0)
JS	−4.7 (0.6)	−2.3 (0.6)
JA	−3.7 (0.6)	−2.0 (0.0)
RZ	−5.3 (0.6)	−7.3 (0.6)
JB	−7.0 (1.0)	−4.7 (1.2)
TD	−13.3 (1.5)	−3.7 (0.6)
SP	−9.0 (1.0)	−3.0 (1.0)

B. Results

Signal levels at threshold, relative to the pedestal level (as in the upper panels of Fig. 3), are given in Table I for the 120- and 180-degree conditions. For three of the four original listeners (AO, JS, and JA), thresholds for the 120-degree condition are higher than for the 180-degree (decrement) condition. This difference is significant for all three listeners $[t(4) > 4.9; \ p < 0.01]$, and suggests that spectral splatter was not detected in the decrement conditions of experiment 1. However, for listener RZ, thresholds for the 120-degree condition are *lower* by 2 dB. This difference in the opposite direction is also significant $[t(4) = 4.25; \ p < 0.05]$. This exception is puzzling. If listener RZ could detect the spectral splatter, both conditions should be equally detectable. However, there seems to be no ready explanation for why the 120 degree should be more detectable. Also, RZ seems much more sensitive to detection in the 120-degree condition than the other three listeners. This unexpected result led to the experiment being repeated using three new listeners, all of whom had extensive experience in psychoacoustic tasks. Their results are presented in the lower three rows of Table I. All show the same (significant) trend as three of the four original listeners: the 180-degree condition produces lower thresholds than the 120-degree condition. Listener TD was considerably more sensitive than any other listener in the 180-degree condition. Note also that while there were no measurable practice effects for the listeners of experiment 1, the more experienced listeners (AO, JB, TD, and SP) generally achieved lower thresholds in the 180-degree condition than the less experienced listeners (JA, JS, and RZ).

The results indicate that, for three of the four original listeners and for all three new listeners, spectral splatter was probably not detectable for the stimuli of experiment 1, even in the absence of noise. However, the difference in thresholds between the two conditions was generally small, and one listener (RZ) showed the opposite effect. Auditory filters are thought to widen with increasing center frequency. Therefore, spectral splatter should become less detectable with increasing pedestal frequency. In the following experiment increment and decrement thresholds were measured using a pedestal and signal frequency of 8 kHz. At this frequency, it was possible to omit the background noise without any risk of introducing audible spectral splatter. Temporal resolution is thought to be approximately independent of center frequency, at least above about 1 kHz (Moore *et al.*,
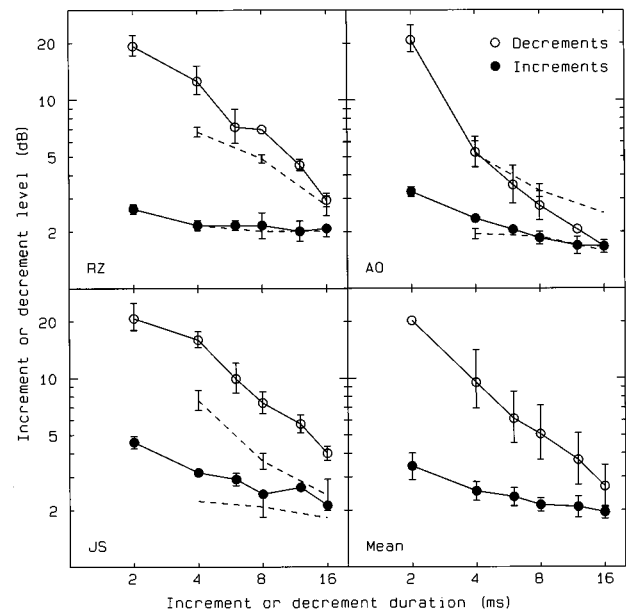
FIG. 6. Individual and mean thresholds for increment and decrement detection with an 8-kHz, 55 dB SPL sinusoidal pedestal. Dashed lines show thresholds with the 4-kHz sinusoidal pedestal, replotted from Fig. 1. Error bars represent ±1 standard error of the mean.

1993b). Thus it was expected that the results at 8 kHz should be very similar to those at 4 kHz.

IV. EXPERIMENT 4. INCREMENT AND DECREMENT DETECTION AT 8 kHz

A. Method

Thresholds for increments and decrements were measured in the absence of noise at a frequency of 8 kHz. Again, the stimuli and procedure were the same as in experiment 1, with the following exceptions: thresholds were measured for half-amplitude signal durations of 2, 4, 6, 8, 12, and 16 ms; the initial step size of the adaptive procedure was 4 dB and the minimum step size was 1 dB. The minimum step size was reduced in order to allow more accurate estimates of decrement thresholds at very short signal durations. Three listeners (AO, JS, and RZ) participated in the experiment.

B. Results

Individual and mean thresholds, in terms of ΔL, are shown in Fig. 6. Dashed lines show the corresponding thresholds at 4 kHz. As with the results at 4 kHz, there is a large asymmetry between increment and decrement detection. Again, the results of AO are somewhat different from those of JS and RZ. While AO shows little or no effect of pedestal frequency, JS and RZ generally require larger increments and decrements for detection at 8 kHz than at 4 kHz. The reason for this is not clear; few other studies of temporal resolution have included frequencies as high as 8 kHz. The deterioration may be related to the deterioration of intensity discrimination observed at very high frequencies if slow amplitude modulation is used as a measure (Riesz, 1928): an increase in the intensity difference limen would result in a roughly parallel upward shift in thresholds, as is observed for

JS and RZ. The following section examines the results both from 4 and 8 kHz in light of current models of temporal resolution.

V. PREDICTIONS USING THE TEMPORAL-WINDOW MODEL

A. Model structure

The temporal-window model used here is very similar to that used in many previous studies (Moore *et al.*, 1993b, 1996; Oxenham and Moore, 1994; Peters *et al.*, 1995). It is assumed that stimuli are band-pass filtered (simulating the auditory filters), rectified, compressed, and passed through a sliding temporal integrator (temporal window). A decrement or increment in level is ''detected'' if the decrease or increase in the output of the temporal window reaches a certain criterion value, ΔO (in dB).

All stimuli were represented by their temporal envelopes. When included, the background noise was represented by a flat temporal envelope with a level equal to the effective level of the noise within the auditory filter centered at the pedestal frequency. As in previous studies, the noise and pedestal were assumed to add in a way equivalent to quadrature phase. Initially, the stimuli were compressed by raising the envelopes to the power n, where $n=0.7$, as in Peters *et al.* (1995) and Moore *et al.* (1996).

The shape of the temporal window was defined by a pair of exponential functions, described by

$$W(t) = \exp(t/T_b), \quad t \leqslant 0 \tag{1}$$

and

$$W(t) = \exp(-t/T_a), \quad t > 0, \tag{2}$$

where t is time, and T_b and T_a are the time constants determining the sharpness of the function for times before and after the peak, respectively (all in ms). For convenience, the equivalent rectangular duration (ERD) of the function is defined as $T_b + T_a$, although this is strictly only true for the case where $n=2$. The value of T_b was assumed to be 1.5 times the values of T_a, reflecting the fact that the decay of backward masking is steeper than that of forward masking (Oxenham and Moore, 1994). Thus the time constants provide the model with one free parameter, the ERD.

The second free parameter is the value of the decision criterion, ΔO which is defined as $10 \log(S/D)$ for decrements and as $10 \log(P/S)$ for increments, where D is the minimum value at the output of the temporal integrator in response to a decrement, P is the maximum response to an increment, and S is the response during the steady-state portion of the stimulus (pedestal alone). The units (10 log) were chosen to conform with previous studies (e.g., Moore *et al.*, 1993b). The best-fitting values for ΔO and the ERD were estimated from a given data set as described in Moore *et al.* (1993b) and Peters *et al.* (1995).

B. Model predictions in the absence of background noise

The results from experiment 3 indicate that at 4 kHz, spectral splatter was probably not audible in the absence of

TABLE II. Best-fitting model parameters for the individual and mean data in the no-noise condition at 4 kHz. ERDs are given in ms.

Listener	Decrements		Increments	
	ERD	ΔO	ERD	ΔO
AO	11.1	0.62	2.0	0.60
JS	5.4	0.84	2.2	0.67
JA	16.2	0.49	3.7	0.58
RZ	7.9	0.87	1.4	0.70
Mean	8.9	0.72	2.3	0.64

background noise, for at least three of the four listeners. For this reason, only the data from the no-noise conditions were examined initially. Data from increment- and decrement-detection conditions were fitted separately. As in the study of Moore *et al.* (1993b), the longest-duration signals (in this case 200 ms) were not included in the fit, as detection in these conditions may be enhanced by a ''multiple-looks'' strategy (Viemeister and Wakefield, 1991) not accounted for within this model. Both individual data and mean data were fitted.

Table II shows the resulting parameters from the 4-kHz pedestal conditions. As noted previously (Buus and Florentine, 1985; Peters *et al.*, 1995; Moore *et al.*, 1996), there is a trade-off between the values of the ERD and ΔO. Thus a large value for the ERD with a small ΔO can sometimes produce similar results to a small ERD with a large ΔO. Something of this relationship can be seen in the best-fitting values for the decrement data (left two columns of Table II). For instance, the largest value of the ERD (16.2 ms) is found for listener JA, together with the smallest value of ΔO (0.49). This trade-off can account for much of the individual differences in parameters for decrement detection. The ERD of 8.9 ms for the mean data is in good agreement with previous estimates of the temporal-window shape both from nonsimultaneous masking and decrement-detection experiments (Oxenham and Moore, 1994; Peters *et al.*, 1995). In contrast, the best-fitting ERDs for the increment conditions (right two columns) are much smaller than those for the decrement conditions. This is true for all four listeners and the mean data. The difference cannot be accounted for in terms of the trade-off between the ERD and ΔO. In fact, for three of the four listeners, ΔO is also smaller in the increment conditions. For the mean data, the ERD is smaller for increments than for decrements by a factor of nearly 4.

A similar pattern of results can be seen for the 8-kHz data in Table III. Again, estimated ERDs are much greater

TABLE III. Best-fitting model parameters for the individual and mean data at 8 kHz.

Listener	Decrements		Increments	
	ERD	ΔO	ERD	ΔO
AO	18.0	0.33	3.0	0.59
JS	6.9	1.42	3.1	0.82
RZ	8.5	1.04	1.2	0.73
Mean	8.56	0.88	2.34	0.72
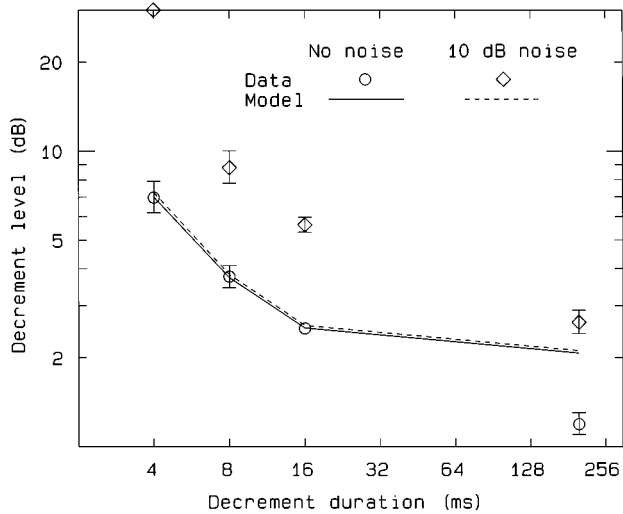
FIG. 7. Effects of a 10-dB spectrum level background noise on the predictions of the temporal-window model (curves), compared with the mean results from experiment 1 (symbols).

for decrement than for increment conditions. Thus in contrast to the assumptions of the temporal-window model, increment- and decrement-detection data produce very different estimates of temporal resolution.

C. Predicted effects of background noise

Within the temporal-window model, the background noise has generally been simulated by a sinusoid, added in quadrature phase to the pedestal, at a level equal to the effective level of the noise falling within the passband of the auditory filter centered at the pedestal frequency. The effect of the noise simulation on model predictions was compared with the effect of noise on listeners' performance in decrement detection. Using the best-fitting parameters for the mean 4-kHz data (see Table II), thresholds in the decrement-detection task were predicted both without noise and with the effect of a 10-dB spectrum level noise taken into account. A 10-dB spectrum level noise has an overall level within the ERB around 4 kHz of 36.6 dB SPL. The resulting predictions, together with the relevant mean data (taken from Fig. 3) are shown in Fig. 7. Thresholds in the absence of noise are predicted well (solid curve). This is to be expected, as the parameters were derived from this data set (excluding the 200-ms point). However, as shown by the dashed curve, predictions are hardly affected by the addition of the simulated 10-dB spectrum level noise. This contrasts strongly with the data which show a marked deterioration due to the noise. Thus the temporal-window model fails to account for the effects of background noise.

In a second approach, mean decrement thresholds in the presence of different noise levels were fitted independently with the temporal-window model. Again, signal durations of 4, 8, and 16 ms were used. For the purposes of modeling it was assumed that listeners could just detect a complete 4-ms gap in the 10-dB noise decrement condition. The best-fitting parameters at each noise level are given in Table IV. As can be seen, the main change is an increase in the detection criterion, ΔO, with increasing noise level; the ERD stays rea-

TABLE IV. Best-fitting model parameters for the mean decrement data at 4 kHz using background-noise spectrum levels between −20 and 10 dB.

Noise spectrum level	ERD	ΔO
−20 dB	9.3	0.75
−10 dB	8.1	0.92
0 dB	9.3	1.08
10 dB	7.2	1.71

sonably constant. This provides some support for the temporal-window model: the noise may have a distracting effect which increases the minimum detectable change in the output of the temporal integrator, but the properties of the integrator itself are not affected. Nevertheless, no quantitative way of accounting for the effects of noise is currently available.

D. Influence of compression on predicted thresholds

While the temporal-window model cannot accurately predict both increment and decrement detection using the same parameters, *some* of the asymmetry apparent in the data is predicted. The size of the predicted asymmetry depends on the amount of compression used. In the simulations described so far the stimulus envelopes were raised to the power $n=0.7$, which is equivalent to raising stimulus intensity to the power 0.35. This value was also used in previous studies (Moore *et al.*, 1993b, 1996; Peters *et al.*, 1995), but it is somewhat less compressive than most psychophysical estimates of peripheral compression, which have centered around $n=0.4$ (e.g., Oxenham and Moore, 1994, 1995; Oxenham and Plack, 1997; Oxenham *et al.*, 1997). On the other hand, earlier work on temporal resolution often operated on (uncompressed) amplitude ($n=1$; Viemeister, 1979; Forrest and Green, 1987) or intensity-like ($n=2$; Moore *et al.*, 1988; Plack and Moore, 1990) quantities.

Simulations showed that reducing the value of n (i.e., increasing the amount of compression) decreased the predicted asymmetry, making predictions worse. Eliminating compression and integrating an intensity-like quantity ($n=2$) increased the predicted asymmetry, but not sufficiently to match the data. This is illustrated in Table V. There, the best-fitting parameters for the mean data at 4 kHz are given using values of n of 0.3 and 2. While the ERDs for the increments and decrements are less disparate using $n=2$, they still differ by more than a factor of 2.

In summary, using a different compression exponent cannot resolve the discrepancy between the data and the model predictions. In fact, a more ''realistic'' compression exponent than was initially used produces a greater discrepancy.

TABLE V. Best-fitting model parameters for the mean data at 4 kHz, using compression exponents of $n=0.3$ and $n=2$.

	Decrements		Increments	
n	ERD	ΔO	ERD	ΔO
0.3	10.7	0.29	2.2	0.27
2.0	5.8	3.45	2.5	1.82
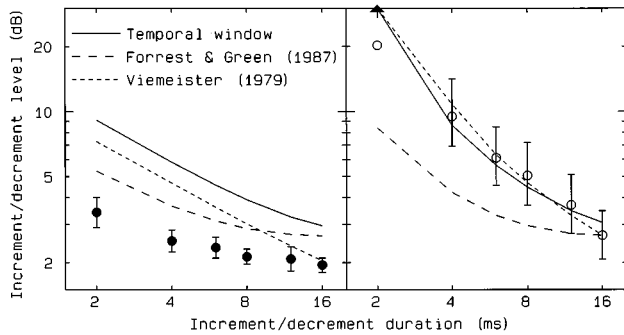
FIG. 8. Predictions of three models of temporal resolution, compared with the mean data from experiment 4, using an 8-kHz pedestal. Increment and decrement thresholds are shown in the left and right panels, respectively. The upward-pointing arrow in the right panel represents a decrement duration at which the model simulations joined to the arrow predicted that a complete gap was undetectable.

VI. ALTERNATIVE MODELS OF TEMPORAL RESOLUTION

In all the following simulations the mean 8-kHz data are used as the reference. It seems certain that spectral splatter did not play a role at 8 kHz and, furthermore, thresholds were measured for a larger number of short durations at 8 kHz than at 4 kHz.

A. Two other models from the literature

Viemeister (1979) proposed a model of temporal resolution to account for the detection of sinusoidal amplitude modulation in a broadband noise carrier. This model comprises the following stages: a bandpass ''pre-detection'' filter; a half-wave rectifier; a first-order low-pass filter; and a decision device based on the standard deviation of the output of the low-pass filter. A similar model was used by Forrest and Green (1987) to account for both gap detection and modulation detection in a broadband noise. The primary difference between the two models lies in the detection device, which for Forrest and Green is based on the ratio between the maximum and minimum output of the low-pass filter within the observation interval. This is referred to as the max/min decision device and is very similar to the ΔO of the temporal-window model. The pre-detection filter was omitted from the present simulations, as the bandwidth of the stimuli was much less than the 2000–4000 Hz bandwidth usually assumed for that filter. As in Forrest and Green (1987), the low-pass filter had a time constant of 3 ms (cutoff frequency=53 Hz).

First, the max/min decision device was used. The criterion max/min ratio of 2.65 dB was chosen to correctly predict the mean decrement threshold for the 16-ms decrement. Predictions are shown as long-dashed curves in Fig. 8. Next, the model of Viemeister (1979) was tested. Using the standard deviation as the decision statistic presents a potential problem, in that the predictions are dependent on the total duration of the analysis window. However, as long as the window is more than twice the duration of the longest signal, predictions are at least monotonic with increasing signal duration. The analysis window duration was set to 400 ms,

which was the length of the steady-state portion of the pedestal in the experiments. Again, the ''threshold'' value of the standard deviation was set to correctly predict the mean experimental threshold for the 16-ms decrement. The predictions of this model are shown as short-dashed curves in Fig. 8. For comparison, the predictions of the temporal-window using the best-fitting parameters derived from the 8-kHz decrement data (see Table III), are shown as solid curves.

Predictions from all three models fail to mirror the asymmetry observed in the data. Also, compressing the signal prior to the low-pass filtering has the same effect as for the temporal-window model.

B. Smoothed onset detection

The failure of the above models to account satisfactorily for the data may be due to inappropriate preprocessing, an inappropriate decision device, or both. One early approach was to include an approximation of neural adaptation in the preprocessing stages (Irwin and Kemp, 1976; Irwin and Purdy, 1982). While this approach was fairly successful, some important assumptions of the model are open to question. For instance, it is assumed that the onset response to an increment is 2.5 times higher than the steady-state response, regardless of prior stimulation by the pedestal. This is clearly not the case physiologically (Smith and Zwislocki, 1975). Also contrary to available data (e.g., Smith *et al.*, 1985), adaptation was assumed not to affect the response to decrements.

Another approach is to alter the decision device. A number of workers have suggested that in some circumstances *changes* (i.e., rapid onsets or offsets) may provide a more salient cue than overall level (Macmillan, 1971, 1973; Laming, 1986; Oxenham, 1997). As mentioned in the Introduction, Macmillan measured the detectability and discriminability of increments and decrements. The shortest signal duration tested by Macmillan was 15 ms and consequently little or no asymmetry was observed between increment and decrement detection. Also, a simple change detector based, say, on the absolute value of the first derivative of the envelope would not be able to predict the asymmetry observed in the present data. In fact, the present data could be interpreted as indicating that the auditory system is more sensitive to onsets than offsets. Brief increments may be detected by virtue of their onsets, which are independent of increment duration. Brief decrements may be detected by the positive-going slope at the end of the decrement. At very short decrement durations, this slope may be ''masked'' by the prior pedestal, leading to a stronger dependence of thresholds on decrement duration.

This hypothesis was evaluated using a variant of the temporal-window model. The only change to the model was that the decision device based on ΔO was replaced by a decision device based on the maximum (positive) *slope* of the temporal-window output. It was therefore assumed that detection occurred when the slope of the output exceeded a certain criterion value, relative to the steady-state output of the temporal window. Simulations of the 8-kHz data were run using exponents of $n=0.7$ (compression) and $n=2.0$ (intensity-like processing). The model parameters (ERD and
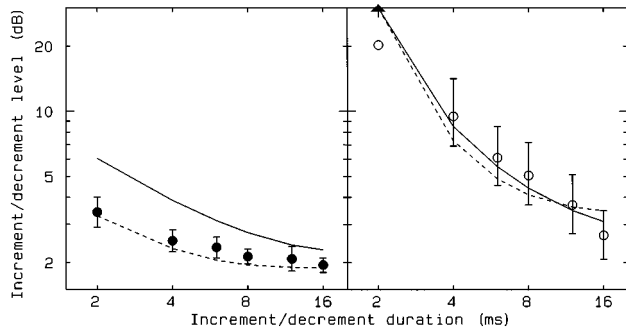
FIG. 9. Predictions of the temporal-window model, using a ''maximum slope'' decision device, compared with the mean data from experiment 4. Increment and decrement thresholds are shown in the left and right panels, respectively. The solid curves represent predictions using an exponent $n = 0.7$ (compression), and the dashed curves represent predictions using an exponent $n = 2.0$ (intensity-like processing). The upward-pointing arrow in the right panel denotes a decrement duration at which both models predicted that a complete gap would not be detectable.

criterion slope value) were derived using data from the 8-kHz decrement-detection condition. The resulting ERDs were 12.1 and 5.9 ms for the $n = 0.7$ and $n = 2$ conditions, respectively. The results from these simulations are shown in Fig. 9. The predictions using the compressive nonlinearity (solid curves) still overestimate thresholds in the increment condition, although the difference is less than for the original decision device (see solid curve in Fig. 8). The predictions using the square-law device are better, and generally lie within one standard error of the mean data.

A growing body of evidence suggests that it is important to include a compressive nonlinearity in models of auditory processing (Penner, 1980; Penner and Shiffrin, 1980; Oxenham and Moore, 1994, 1995; Oxenham and Plack, 1997; Moore and Oxenham, 1997). It is therefore somewhat surprising that the best fit was achieved using a square law, rather than a compressive nonlinearity. Nevertheless, the simulations show that a straightforward change in the decision criterion can improve the fit considerably for both types of nonlinearity.

The temporal-window model has been successfully used to account for a fairly wide range of psychoacoustic data, including the temporal decay of forward and backward masking (Moore *et al.*, 1988), the additivity of nonsimultaneous masking (Oxenham and Moore, 1994, 1995), decrement detection (Plack and Moore, 1991) and gap detection (Moore *et al.*, 1993a). Figure 9 suggests that modifying the decision device may enable the model to also predict at least some aspects of the data presented in this study.

VII. SUMMARY

Thresholds for detecting increments and decrements in the level of a 55-dB sinusoid were measured at 4 kHz as a function of signal duration over a range of background-noise levels. Thresholds were also measured at 8 kHz in the absence of noise. The results were compared with simulations using a temporal-window model and other models of temporal resolution. The following conclusions were reached:

(1) At the shortest durations there was a marked asymmetry between increment and decrement detection, when measured in terms of ΔL or signal level $20 \log(\Delta p/p)$: for a given duration, decrements were less easily detected than increments.

(2) Thresholds were dependent on background-noise level over all noise levels tested. The results are not due to the masking of spectral splatter.

(3) Neither aspect of the data mentioned above could be accounted for by the temporal-window model in its standard form. While the decrement-detection data produced estimates of the ERD comparable to those of previous studies (around 9 ms), the increment-detection data were best described by an ERD of less than 2.5 ms, which is smaller by nearly a factor of 4. Two other popular models of temporal resolution (Viemeister, 1979; Forrest and Green, 1987) also failed to account for the data.

(4) Manipulations of the spectral content of the background noise revealed no simple relationship between noise spectrum and signal thresholds. It is possible that off-frequency noise fluctuations interfere in the detection of fluctuations caused by the signal.

(5) A different decision criterion, based on the maximum positive slope of the temporal-window output, produced a better fit to the data in the absence of noise. This may indicate that positive-going slopes (or onsets) are of more perceptual importance than is attributed to them in current models of temporal processing.

APPENDIX: SIGNAL PHASE AND SPECTRAL SPLATTER

It has been claimed that spectral splatter should be independent of the phase relationship between a long-duration sinusoidal masker and a brief signal of the same frequency (Leshowitz and Wightman, 1971). This is because it is assumed that the masker can be represented by a line spectrum. Thus all off-frequency spectral components of the signal do not interact with the masker. This assertion, important to the interpretation of all the experiments presented here, was tested directly using three listeners. Two listeners (AO and JB) also participated in other experiments in this study; the third (JV) was a student who had some previous experience in psychoacoustic tasks. The measurement procedure was the same as that used in experiment 1. The pedestal was also the same as in experiment 1. The signal was temporally centered in the pedestal, had a total duration of 4 ms, and was gated abruptly. Its starting phase was always zero. The phase relationship between the masker and the signal was the independent variable and was set to 0, 120, and 180 degrees. This was achieved by setting the starting phase of the pedestal to 0, 240, or 180 degrees. As the signal was very brief and was

gated abruptly, it was expected that performance would be limited by the detection of spectral splatter. If, as expected, spectral splatter is independent of the masker-signal phase relationship, then thresholds from all three conditions should not be significantly different from each other.

This prediction was confirmed in the results. Thresholds for each condition and listener were repeated 3 times. The mean thresholds, relative to pedestal level, were -17.9, -17.0, and -17.3 dB for the 0-, 120-, and 180-degree conditions, respectively. A within-subjects analysis of variance confirmed that there was no significant effect of phase [$F(2,4)=1.96$; $p>0.25$]. Thus it seems reasonable to assume that when splatter is detectable, performance is independent of the phase relationship between signal and pedestal.

---

[1]While of off-frequency energy is independent of the phase relationship between the pedestal and the signal, this is not true for the phase relationship between the signal window and the signal itself. However, in this case, the ramps of the window (2 ms) are large compared with the signal period (0.25 ms), meaning that the change in off-frequency energy with changing window-signal phase is negligible. The signals used in this study did not exceed 55 dB SPL. Thus any part of the signal spectrum which was more than 55 dB below the peak was certainly not detectable. (In fact, absolute thresholds for a 4-ms signal rarely fall below 20 dB SPL.) Using a 55-dB dynamic range for analysis, the maximum difference in the spectra of the 4-ms signals with 0 and 120 degree starting phases was less than 0.06 dB.

Bacon, S. P., and Grantham, D. W. (**1989**). ''Modulation masking: Effects of modulation frequency, depth and phase,'' J. Acoust. Soc. Am. **85**, 2575–2580.

Buunen, T. J. F., and van Valkenburg, D. A. (**1979**). ''Auditory detection of a single gap in noise,'' J. Acoust. Soc. Am. **65**, 534–537(L).

Buus, S., and Florentine, M. (**1985**). ''Gap detection in normal and impaired listeners: The effect of level and frequency,'' in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, New York), pp. 159–179.

Buus, S., and Florentine, M. (**1991**). ''Psychometric functions for level discrimination,'' J. Acoust. Soc. Am. **90**, 1371–1380.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). ''Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers,'' J. Acoust. Soc. Am. (submitted).

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). ''Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,'' J. Acoust. Soc. Am. (submitted).

de Boer, E. (**1986**). ''On thresholds of short-duration intensity increments and decrements,'' in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson, (Plenum, New York), pp. 429–436.

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.

Gerken, G. M., Bhat, V. K. H., and Hutchinson-Clutter, M. (**1990**). ''Auditory temporal integration and the power function model,'' J. Acoust. Soc. Am. **88**, 767–778.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hearing Res. **47**, 103–138.

Houtgast, T. (**1989**). ''Frequency selectivity in amplitude-modulation detection,'' J. Acoust. Soc. Am. **85**, 1676–1680.

Irwin, R. J., and Kemp, S. (**1976**). ''Temporal summation and decay in hearing,'' J. Acoust. Soc. Am. **59**, 920–925.

Irwin, R. J., and Purdy, S. C. (**1982**). ''The minimum detectable duration of auditory signals for normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **71**, 967–974.

Laming, D. (**1986**). *Sensory Analysis* (Academic, London).

Leshowitz, B., and Wightman, F. L. (**1971**). ''On-frequency masking with continuous sinusoids,'' J. Acoust. Soc. Am. **41**, 1180–1190.

Leshowitz, B., and Wightman, F. L. (**1972**). ''On the importance of considering the signal's frequency spectrum: Some comments on Macmillan's 'Detection and recognition of increments and decrements in auditory intensity' experiment,'' Percept. Psychophys. **12**, 209–210.

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.

Mendoza, L., Hall, J. W., and Grose, J. H. (**1995**). ''Modulation detection interference using random and sinusoidal amplitude modulation,'' J. Acoust. Soc. Am. **97**, 2487–2497.

Moore, B. C. J., and Oxenham, A. J. (**1997**). ''Psychoacoustic consequences of compression in the peripheral auditory system,'' Psychol. Rev. (in press).

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1993a**). ''Detection of temporal gaps in sinusoids: Effects of frequency and level,'' J. Acoust. Soc. Am. **93**, 1563–1570.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1993b**). ''Effects of frequency on the detection of decrements and increments in sinusoids,'' J. Acoust. Soc. Am. **94**, 3190–3198.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1996**). ''Detection of decrements and increments in sinusoids at high overall levels,'' J. Acoust. Soc. Am. **99**, 3669–3677.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (**1988**). ''The shape of the ear's temporal window,'' J. Acoust. Soc. Am. **83**, 1102–1116.

Oxenham, A. J. (**1997**). ''Temporal integration at 6 kHz as a function of masker bandwidth,'' J. Acoust. Soc. Am. (submitted).

Oxenham, A. J., and Moore, B. C. J. (**1994**). ''Modeling the additivity of nonsimultaneous masking,'' Hearing Res. **80**, 105–118.

Oxenham, A. J., and Moore, B. C. J. (**1995**). ''Additivity of masking in normally hearing and hearing-impaired subjects,'' J. Acoust. Soc. Am. **98**, 1921–1934.

Oxenham, A. J., and Plack, C. J. (**1997**). ''A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing,'' J. Acoust. Soc. Am. **101**, 3666–3675.

Oxenham, A. J., Moore, B. C. J., and Vickers, D. A. (**1997**). ''Short-term temporal integration: Evidence for the influence of peripheral compression,'' J. Acoust. Soc. Am. **101**, 3676–3687.

Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (**1992**). ''Complex sounds and auditory images,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford).

Penner, M. J. (**1980**). ''The coding of intensity and the interaction of forward and backward masking,'' J. Acoust. Soc. Am. **67**, 608–616.

Penner, M. J., and Shiffrin, R. M. (**1980**). ''Nonlinearities in the coding of intensity within the context of a temporal summation model,'' J. Acoust. Soc. Am. **67**, 617–627.

Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (**1995**). ''Effects of level and frequency on the detection of decrements and increments in sinusoids,'' J. Acoust. Soc. Am. **97**, 3791–3799.

Plack, C. J., and Moore, B. C. J. (**1990**). ''Temporal window shape as a function of frequency and level,'' J. Acoust. Soc. Am. **87**, 2178–2187.

Plack, C. J., and Moore, B. C. J. (**1991**). ''Decrement detection in normal and impaired ears,'' J. Acoust. Soc. Am. **90**, 3069–3076.

Riesz, R. R. (**1928**). ''Differential intensity sensitivity of the ear for pure tones,'' Phys. Rev. **31**, 867–875 .

Smith, R. L. (**1979**). ''Adaptation, saturation, and physiological masking in single auditory-nerve fibers,'' J. Acoust. Soc. Am. **65**, 166–178.

Smith, R. L., Brachman, M. L., and Frisina, R. D. (**1985**). ''Sensitivity of auditory-nerve fibers to changes in intensity: A dichotomy between decrements and increments,'' J. Acoust. Soc. Am. **78**, 1310–1316.

Smith, R. L., and Zwislocki, J. J. (**1975**). ''Short-term adaptation and incremental responses in single auditory-nerve fibers,'' Biol. Cybern. **17**, 169–182.

Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based on modulation thresholds,'' J. Acoust. Soc. Am. **66**, 1364–1380.

Viemeister, N. F., and Wakefield, G. H. (**1991**). ''Temporal integration and multiple looks,'' J. Acoust. Soc. Am. **90**, 858–865.

Yost, W. A., and Sheft, S. (**1989**). ''Across-critical-band processing of amplitude-modulated tones,'' J. Acoust. Soc. Am. **85**, 848–857.

Zwicker, E. (**1956**). ''Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs,'' Acustica **6**, 356–381.

Binaural modulation detection interference

Stanley Sheft and William A. Yost

Parmly Hearing Institute, Loyola University of Chicago, 6525 N. Sheridan Road, Chicago, Illinois 60626

The ability to detect amplitude modulation (AM) of a tonal probe can be disrupted by the presence of modulated masking tones. Two experiments examined whether a disparity in the interaural parameters of the probe and masker can reduce the amount of interference. In the first experiment, the effects of interaural time and intensity differences were studied in separate sets of conditions. With low-frequency carriers, the detection of 10-Hz probe modulation in the presence of 10-Hz masker modulation was not significantly affected by interaural time differences. With higher-frequency carriers, dichotic stimuli were generated through combinations of diotic, dichotic, or monotic probe and masker presentations in which the probe and masker did not share a common interaural intensity difference. In these conditions, the amount of interference was affected by the interaural configuration. However, monotic level differences between the probe and masker may have contributed to the effect of interaural configuration. In the second experiment, the probe and masker were presented through separate speakers in an enclosed listening environment. Spatial separation between the sources for the probe and masker led to a small reduction in the amount of interference. When the masker modulation rate was varied with the probe AM rate fixed at 10 Hz, the extent of tuning in the modulation domain in the sound-field conditions was similar to that obtained with diotic stimulus presentation over headphones. © *1997 Acoustical Society of America.* [S0001-4966(97)03109-3]

PACS numbers: 43.66.Mk, 43.66.Pn [WJ]

INTRODUCTION

Modulation detection interference (MDI) refers to the elevation in threshold for processing modulation of a tonal carrier due to the presence of modulation in another spectral region. Interference has been observed in tasks involving amplitude modulation (AM) detection (Yost and Sheft, 1989, 1990, 1994; Hall and Grose, 1991; Bacon and Konrad, 1993; Bacon and Moore, 1993; Bacon and Opie, 1994; Bacon *et al.*, 1995; Mendoza *et al.*, 1995), AM depth discrimination (Moore *et al.*, 1991; Moore and Shailer, 1992, 1994; Shailer and Moore, 1993; Yost and Sheft, 1994), and AM rate discrimination (Yost *et al.*, 1989). Due to the presence of MDI with wide separations between the probe and masker carrier frequencies, Yost and Sheft (1989) suggested that the interference is a consequence of processing that uses common modulation to group spectral components. However, results from studies in which the probe and masker modulators did not share either a common rate, phase, or pattern show greater interference than might be expected from grouping effects based on coherent modulation (Yost and Sheft, 1989, 1994; Yost *et al.*, 1989; Moore *et al.*, 1991; Shailer and Moore, 1993; Bacon and Konrad, 1993; Bacon and Opie, 1994; Bacon *et al.*, 1995). In these studies, MDI was not always greatest when the probe and masker modulations were identical. These results indicate either a relatively broad selectivity to modulation characteristics in the grouping process, involvement of other factors as the source of the interference, or some combination of these two alternatives. Moore and Shailer (1992) proposed two sources of the interference in MDI, with only one sensitive to perceptual grouping of the probe and masker. They termed the component that relates to grouping as carrier-specific MDI, suggesting

that it arises from difficulty in "hearing out" the probe from the masker. The other component, termed modulation-specific MDI, was attributed to difficulty in distinguishing the modulation of the probe from the masker modulation. Modulation-specific MDI could then persist in conditions in which the probe and masker carriers can be heard out as separate entities.

Several studies have evaluated the relationship between MDI and grouping by controlling stimulus factors intended to affect the perceptual grouping of the probe and masker carriers. The common assumption of these studies was that MDI would be reduced in conditions in which stimulus characteristics promote perceptual segregation of the probe and masker. Factors studied include stimulus gating, duration, harmonicity, and cuing of the probe carrier frequency. Hall and Grose (1991) showed that the amount of MDI gradually decreased with increasing the gating asynchrony between the probe and masker carriers. Subsequent studies have demonstrated a substantial reduction in MDI with just an onset or offset gating asynchrony (Moore and Shailer, 1992) or with a silent gap during the probe carrier presentation (Sheft, 1997). With synchronous gating of the probe and masker, increasing stimulus duration from 200 to 1600 ms had little effect on MDI (Moore and Shailer, 1992). The presence or absence of a harmonic relationship between the probe and masker carriers was also not a strong factor in MDI (Bacon and Moore, 1993; Moore and Shailer, 1994). Moore and Shailer (1992) reported a small reduction in MDI in an AM depth discrimination task when an unmodulated tone burst was used to cue the probe carrier frequency before each trial. In an AM detection task, Sheft (1997) found that cuing led to a slight reduction in MDI for some subjects, while for others it had

no effect. Among these four stimulus factors designed to promote segregation of the probe and masker, only asynchronous gating of the probe and masker carriers had a large effect.

Perceptual grouping can be influenced by the coherence of binaural information. Interaural phase or level differences can segregate a narrow-band from a wideband stimulus (e.g., Cramer and Huggins, 1958; Yost, 1991a) or a single component from a tonal complex (Kubovy and Daniel, 1983). Yost and Sheft (1990), Bacon and Opie (1994), and Mendoza *et al.* (1995) measured MDI with the probe presented to one ear and the masker to the other. MDI was obtained in these studies with this dichotic mode of stimulus presentation. The amount of interference with dichotic stimulus presentation was less than that obtained in the diotic or monotic reference condition. These results are consistent with a reduction in MDI due to segregation of the probe and masker by laterality. This consideration, however, is tempered by several concerns. Across the three studies cited above, the reduction in the amount of MDI between the dichotic and either diotic or monotic reference condition varied from an average value of less than 1 dB in the work of Mendoza *et al.* (1995) to over 7 dB in the study by Yost and Sheft (1990). Also, these studies used only the one dichotic stimulus configuration in which the probe and masker were presented to opposite ears. Finally, if peripheral interaction of the probe and masker accounted for any of the interference in the diotic or monotic reference condition of these studies, the elimination of this interaction with the probe and masker presented to opposite ears would result in a reduction in MDI unrelated to binaural processing per se.

The purpose of the present work was to further study the relationship between MDI and auditory grouping by determining the extent to which probe and masker segregation based on binaural cues can reduce the amount of interference. MDI was measured in a variety of dichotic listening conditions, many in which the probe and masker were still mixed together in one or both ears. In the first experiment, stimuli were presented over headphones. Conditions evaluated the effects of interaural time differences (ITDs) with low-frequency carriers for the probe and masker and interaural intensity differences (IIDs) with high-frequency carriers. In the second experiment, the probe and masker were presented through separate speakers in an enclosed sound field with the spatial separation between the two sources varied across conditions.

I. EXPERIMENT 1: LATERALIZATION CONDITIONS

A. Method

1. Conditions and stimuli

The task was to detect AM of a probe tone. The probe AM rate was 10 Hz. The probe was either presented alone or in the presence of a two-carrier masker complex. The masker was either not modulated or modulated at 10 Hz with the probe and masker modulation in phase. When modulated, the masker modulation index was 1.0. Except in two of the IID conditions, the level of the probe tone and of each masker carrier was 65 dB SPL. In the two IID conditions, the level at one ear of either the masker or of both the probe and masker was lowered by 6 dB (see below). The rms pressure of the modulated carriers was held constant by dividing the modulated waveform by $(1 + m^2/2)^{1/2}$ where m is the modulation index. Stimuli were 400 ms in duration and were shaped by 10-ms $\cos^2$ rise–fall ramps.

The effects of ITDs and IIDs on MDI were evaluated in separate sets of conditions. In the ITD conditions, the probe carrier frequency was 500 Hz and the carrier frequencies of the masker complex were 223 and 1038 Hz. ITDs were varied by changing the carrier phase arguments. Three stimulus configurations were used: (1) the probe and masker carriers were diotic, (2) the probe carrier was generated with a 90° interaural phase difference (IPD) leading to the right ear while the masker remained diotic, and (3) the probe carrier had a 90° IPD leading to the right ear and each masker carrier had a 90° IPD leading to the left ear. The probe and masker modulators were always interaurally in phase.

In the IID conditions, the probe carrier frequency was 2049 Hz and the carrier frequencies of the masker complex were 909 and 4186 Hz. Five binaural stimulus configurations were used: (1) the probe and masker were both presented diotically, (2) the probe was presented to just the right ear and the masker was diotic, (3) the probe was diotic and the masker was presented to just the left ear, (4) the probe was presented to the right ear and the masker to the left ear, and (5) the probe and masker were both presented binaurally with the level of the probe lowered by 6 dB in the left ear and the masker level lowered by 6 dB in the right ear. As a control for the fifth condition, one monotic condition was run in which the probe and masker were presented to only the right ear with the masker level lowered by 6 dB.

A Masscomp minicomputer interfaced with a 16-bit digital-to-analog converter was used for stimulus generation and experimental control. Stimuli were generated at a 20-kHz sampling rate. Following analog conversion, stimuli were low-pass filtered (Rockland model 2382) at 8 kHz. Stimuli were presented through Sony MDR-V6 headphones with the subjects seated in a double-walled soundproof booth.

2. Procedure

Thresholds were measured using a two-interval two-alternative forced-choice (2IFC) tracking procedure (Levitt, 1971). The two observation intervals were separated by 500 ms. Feedback was provided after each trial. The task was to identify the interval in which the probe carrier was modulated. The modulation index (m) for the probe modulation was decreased 2 dB in steps of $20 \log m$ following two consecutive correct responses and increased 2 dB following one incorrect response. A run was terminated after 50 trials. Runs containing less than ten reversals were discarded. Thresholds are reported in terms of $20 \log m$. Excluding the first four reversals (five if the total number of reversals was odd), threshold was estimated as the average of the levels in dB of the remaining reversals. Reported threshold values are the mean of at least six threshold estimates per condition. Sub-
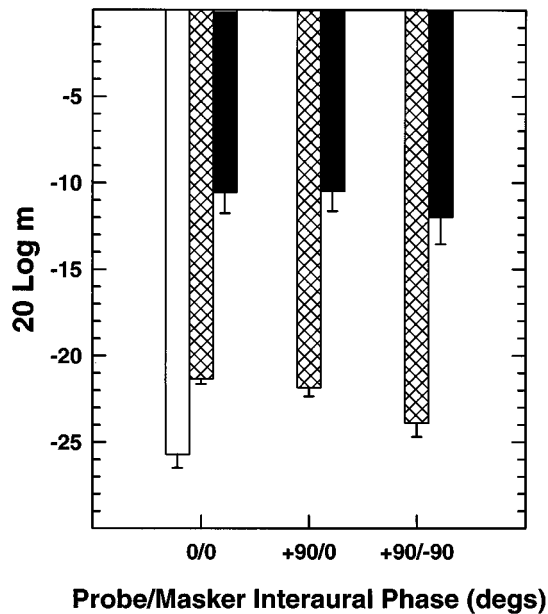
FIG. 1. Thresholds for detecting probe modulation averaged across the five subjects for three configurations of ITDs. The interaural phase of the probe and masker for each configuration is indicated along the abscissa with the labels to the left of the slash referring to the probe and those to the right referring to the masker. The masker was either not present (open bars), not modulated (cross-hatched bars), or modulated (solid bars). Error bars represent the standard error of the mean threshold.

jects were informed of the expected lateral position of the probe and masker in each experimental condition.[1]

3. Subjects

Data were collected from five subjects with pure-tone thresholds with 10 dB of ANSI (1989) standard audiometric thresholds. Subjects received a minimum of 16 h of practice before data collection began.

B. Results and discussion

1. Interaural time differences

In the three stimulus configurations used to evaluate the effect of ITDs on MDI, the probe and masker carriers were each either diotic or presented with an IPD of 90°. If the probe or masker were presented alone, these conditions would contrast an image lateralized at midline to one displaced almost fully towards one ear. Sensitivity to the lateral position of a target complex is often degraded by the addition of distractor tones (e.g., Dye, 1990; Buell and Hafter, 1991; Stellmack and Dye, 1993). For the stimuli used in the first experiment, the contrast between a 0° and 90° IPD for the probe or masker would remain well above threshold when both the probe and masker were present. Results from stimulus classification procedures in which components of a tonal complex had different ITDs suggest that subjects would tend to correctly lateralize the probe to the right of midline in the two dichotic conditions (Dye *et al.*, 1994, 1996; Stellmack and Lutfi, 1996).

Though AM thresholds varied across subjects in each condition of experiment 1, the patterns of results were similar. Consequently, mean data are presented. Figure 1 shows

probe modulation detection thresholds for the three configurations of ITDs. For each configuration, the interaural phase of the probe and masker is indicated along the abscissa with the labels to the left and right of the slash referring to the probe and masker, respectively. The open bar shows the mean threshold for detecting 10-Hz modulation of the probe carrier without the masker present. The cross-hatched bars represent thresholds from conditions in which the masker was present but not modulated. These thresholds were roughly 2–5 dB higher than the probe-alone threshold with performance best when the interaural phase of the probe and masker differed by 180° (i.e., the condition labeled +90/−90). The solid bars show thresholds from conditions in which the masker was modulated. Average thresholds in the three modulated-masker conditions fell in the range of −12 to −10 dB, indicating substantial MDI. As in the unmodulated-masker conditions, the average threshold obtained in the presence of the modulated masker was lowest in condition +90/−90.

A repeated-measures analysis of variance (ANOVA) was performed on factors masker modulation index (0.0 or 1.0) and ITD configuration. There was a significant effect of both modulation index ($F_{1,4} = 90.4$, $p < 0.001$) and ITD ($F_{2,8} = 9.5$, $p = 0.008$). The interaction between the two factors was not significant. *Post-hoc* pairwise comparisons were performed among the three ITD conditions at each level of modulation index. The only significant difference was between the diotic and the +90/−90 conditions with the unmodulated masker. These results indicate that the introduction of ITDs for the probe and masker carriers can have a small effect on probe AM detection threshold in the presence of an unmodulated masker but not in the presence of a modulated masker.

2. Interaural intensity differences

The mean thresholds from the conditions involving IIDs are shown in Fig. 2. The interaural configuration for each set of conditions is indicated along the abscissa with the labels to the left of the slash referring to the probe and those to the right referring to the masker. The labels D, R, and L represent diotic, right-ear-only, and left-ear-only presentation, respectively. The labels d(−6) and R(−6) indicate that a carrier level was lowered by 6 dB to 59 dB SPL. The open bars represent the thresholds obtained without the masker present, the cross-hatched bars represent thresholds in the presence of the unmodulated masker, and the filled bars represent thresholds obtained when the masker was modulated.

Consider first the results to the left of the break mark along the abscissa. In these conditions, the interaural configuration was varied by presenting the probe and masker either diotically or to just one ear. Two probe-alone thresholds are shown, one for diotic and the other for monotic presentation of the probe carrier. There was a small but significant binaural advantage for the detection of probe modulation without the masker present ($t_4 = −10.5$, $p = 0.0005$). The diotic probe-alone threshold serves as a reference for MDI in the two conditions in which the probe was diotic (D/D and D/L) and the monotic threshold is the reference for conditions R/D and R/L. For both the monotic and diotic
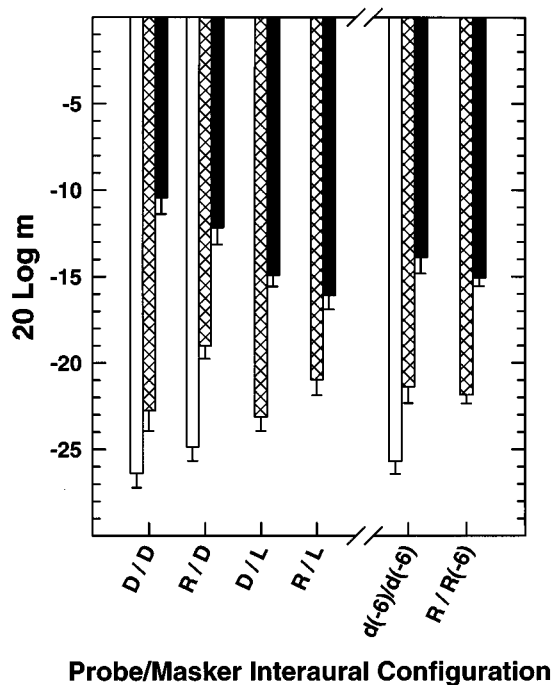
FIG. 2. Thresholds for detecting probe modulation averaged across the five subjects for five configurations of IIDs and one monaural condition. The interaural configuration for each condition is indicated along the abscissa with the labels to the left of the slash referring to the probe and those to the right referring to the masker. The labels D, R, and L represent diotic, right-ear only, and left-ear only presentation, respectively. In condition d(−6)/d(−6), the probe and masker were both presented binaurally with the level of the probe lowered 6 dB in the left ear and the masker level lowered 6 dB in the right ear. In condition R/R(−6), the probe and masker were presented to only the right ear with the masker level lowered by 6 dB. The masker was either not present (open bars), not modulated (cross-hatched bars), or modulated (solid bars). Error bars represent the standard error of the mean threshold.

probe, thresholds rose when the masker was present, with thresholds poorest when the masker was modulated. The pattern of results across the four interaural configurations differed for the unmodulated versus modulated maskers. When the masker was not modulated, thresholds were poorest in condition R/D in which the probe was monotic and the masker was diotic. In this condition, there was also the greatest rise in the mean threshold (5.9 dB) from the appropriate probe-alone reference threshold. In the other three interaural configurations, the rise in the mean threshold due to the addition of the unmodulated masker ranged from 3.3 to 3.9 dB. When the masker was modulated, performance gradually improved across interaural configuration as ordered along the abscissa in Fig. 2, with thresholds lowest when the probe and masker were monotically presented to separate ears (condition R/L). The change in threshold between any two modulated-masker conditions was larger when the masker changed from diotic to monotic (i.e., from D to L) than when the masker's interaural configuration was constant.

A repeated-measures ANOVA on factors masker modulation index (0.0 or 1.0) and IID configuration confirmed these observations. There was a significant effect of both modulation index ($F_{1,4} = 59.3$, $p = 0.002$) and IID ($F_{3,12} = 21.9$, $p < 0.001$). The interaction between the two factors

was also significant ($F_{3,12} = 24.9$, $p < 0.001$) indicating a different pattern of results for the unmodulated versus modulated maskers. *Post-hoc* pairwise comparisons were performed among the four IID conditions at each level of modulation index. For the unmodulated masker, thresholds were significantly higher in condition R/D than in either condition with the diotic probe. For the modulated masker, thresholds were significantly lower in condition D/L or R/L than in either condition D/D or R/D. That is, comparisons were significant between conditions which contrasted two masker configurations. When the same masker IID was used in two conditions (e.g., conditions D/D and R/D), pairwise comparisons were not significant. Results then indicate that the beneficial effect of interaural configuration on probe AM detection in the presence of a modulated masker occurs when the probe is present in one ear without the masker.

In condition d(−6)/d(−6), both the probe and masker were presented binaurally with the level of the probe lowered by 6 dB in the left ear and masker level lowered by 6 dB in the right ear. If presented alone, the 6-dB IID of either the probe or masker would lead to an image lateralized roughly midway between center and the ear receiving the more intense sound (e.g., Yost, 1981). As with ITDs, the perceived laterality of a target based on IIDs can be affected by the IID of a distractor (Dye, 1997). In the present study, the perceived lateral separation between the concurrent probe and masker may then have been less than suggested by work in which there was a single stimulus IID. Thresholds from condition d(−6)/d(−6) are shown to the right of the break mark along the abscissa in Fig. 2. The mean probe-alone and unmodulated-masker thresholds were within 1.5 dB of the thresholds obtained with the diotic probe and masker (condition D/D). When the masker was modulated, the mean threshold was 3.5 dB lower in the dichotic than diotic condition. In condition d(−6)/d(−6), the level of the probe carrier was 6 dB greater than the level of the masker carriers in the right ear. In diotic and monotic conditions, a level difference between the probe and masker carriers can affect the amount of MDI (Yost and Sheft, 1990; Bacon and Konrad, 1993). In condition R/R(−6), both the probe and masker were presented to only the right ear with the level of the masker carriers lowered by 6 dB. Compared to performance in the dichotic conditions which utilized a 6-dB IID, there was little change in either the unmodulated-masker or modulated-masker thresholds. The beneficial effect of an IID on probe AM detection in the presence of a modulated masker may then not relate to a laterality cue as much as to the level difference between the probe and masker at each ear.

Thresholds obtained either without the masker or in the presence of the unmodulated masker can be used as a reference for determining the amount of MDI in dB. MDI in dB is the difference in threshold between the modulated-masker condition and the reference condition. Using the unmodulated-masker thresholds as a reference provides a measure of the amount of interference due to masker modulation. However, this measure obscures changes between conditions in which both the unmodulated- and modulated-masker thresholds vary in a similar fashion (see Bacon and

Moore, 1993). In the five IID conditions, quantifying the amount of interference is further complicated by the small differences among the probe-alone thresholds and the interaction between modulation index and interaural configuration. With the intent of this study being to determine the effect of binaural cues on MDI, the change in the amount of interference across conditions is the key issue. When using either the probe-alone or the unmodulated-masker reference, MDI in dB was greatest in condition D/D and smallest in condition R/L with intermediate amounts of MDI obtained in condition $d(-6)/d(-6)$. Between conditions D/D and R/L, the amount of MDI changed by 7.2 dB when referenced to the probe-alone thresholds and 7.5 dB when referenced to thresholds obtained with the unmodulated masker. An alternative measure is to simply evaluate the change across interaural configuration of the thresholds obtained in the presence of the modulated maskers. The average modulated-masker thresholds vary over a 5.7-dB range across the five IID configurations.

In the Introduction, results from studies which evaluated the effects of asynchronous carrier gating, duration, harmonicity, and cuing on MDI were reviewed. A roughly 6- to 7-dB change in the amount of interference as a function of interaural configuration is larger than the reported effects of either duration, harmonicity, or cuing and close to the reduction in interference obtained with asynchronous gating of the probe and masker carriers. Comparisons across studies need to consider differences in stimulus parameters that can affect MDI, most notably carrier frequencies and modulation characteristics. Sheft (1997) used the same carrier frequencies and modulation rate as used in the IID conditions of the present work to evaluate the effect of carrier asynchrony on MDI. Results averaged across seven subjects showed that a delay of the onset of the probe carrier by 160–360 ms lowered probe AM detection thresholds in the presence of a modulated masker by 7 to 8 dB. Along with comparable reductions in MDI, another similarity between the IID results and the asynchrony study is that, though performance improved with a stimulus configuration intended to promote segregation of the probe and masker, the interference was never eliminated in either study. That is, thresholds in the presence of the modulated masker never equaled those obtained with the unmodulated masker or in the probe-alone condition.

II. EXPERIMENT 2: SOUND-FIELD CONDITIONS

Bregman (1990) and Yost (1991b) have argued that the goal of auditory grouping and segregation is to determine the sources of the sounds present in a multisource listening environment. In the second experiment, the probe and masker were presented through loudspeakers in a listening environment. Except in a control condition, separate speakers were used for the probe and masker presentations. Speaker location was varied in large increments in the horizontal plane so that the probe and masker could always be identified as originating from separate sources when presented through separate speakers. Conditions evaluate the effects of spatial separation and the similarity of source modulation rate on MDI.

A. Method

1. Listening environment

Subjects were seated in an enclosed room 3.5-m long by 2.7-m wide by 2.1-m high. The walls were constructed with sound-deadened office partitions and the ceiling with plywood. All floor, wall, and ceiling surfaces were covered with 3-in.-thick sound-attenuating foam (Azonic AZ 3000). Four speakers (Realistic-Minimus 2.5) were suspended from the ceiling 1.5 m from the subject and at a height of 1.2 m. In the frontal horizontal plane, speaker locations were at $-90°$, $-45°$, $0°$, and $90°$ azimuth.

2. Conditions and stimuli

Results from experiment 1 indicated an effect of interaural stimulus parameters only with the higher frequency carriers for the probe and masker. If this effect was due to either segregation of the probe and masker by laterality or a monaural level difference between the probe and masker carriers, an effect would also be expected in the sound-field conditions. Experiment 2 thus used the carrier frequencies of the IID conditions of experiment 1. The probe carrier frequency was 2049 Hz and the masker was a two-carrier complex with component frequencies of 909 and 4186 Hz. The probe and masker were presented to subjects through either the same or different speakers in the listening room. The level of each carrier was adjusted so that when played through a specific speaker, the sound-pressure level measured with a B&K sound level meter (model 2209) at the approximate position of the subject's head was 65 dB(A). Stimuli were 400 ms in duration and were shaped by 10-ms $\cos^2$ rise–fall ramps.

The task was to detect 10-Hz AM of the probe carrier. The masker was either not present, unmodulated, or modulated with an index of 1.0. In the first set of conditions, the masker modulation rate was also 10 Hz with the probe and masker modulation in phase. The spatial separation between the probe and masker speakers was varied across conditions. In the second set of conditions, the probe and masker speaker locations were fixed with the masker modulation rate varying between 2.5 and 40 Hz. Using these parameters for the probe and masker modulation, data were also collected with stimuli presented diotically over headphones. In the headphone conditions, the level of the probe tone and of each masker carrier was 65 dB SPL. In both the sound-field and headphone conditions, the apparatus for stimulus generation and the method for level compensation due to modulation were the same as in experiment 1.

3. Procedure

Testing in both the sound field and under headphones used the same 2IFC procedure and method of threshold estimation as in experiment 1. When tested in the sound field, subjects were instructed to remain seated facing directly forward at all times and to not point their head towards any sound source. An electronic sensor monitored whether subjects remained seated; head position was not monitored.

1795   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

S. Sheft and W. A. Yost: Binaural MDI   1795

Probe Position: -90°

Masker Lateral Position (degs)
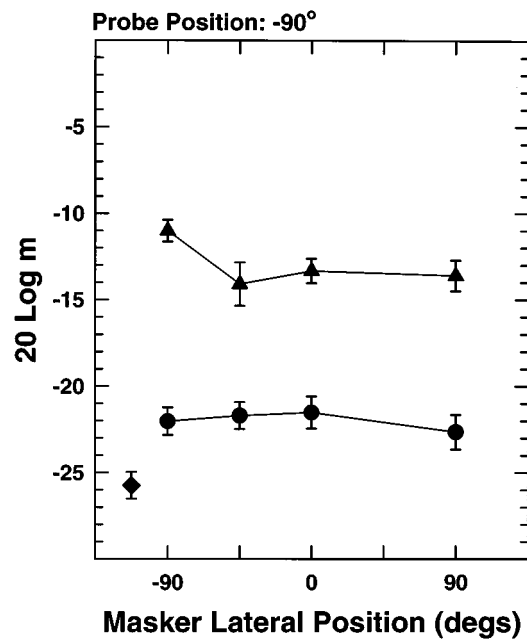
FIG. 3. Thresholds for detecting probe modulation averaged across the six subjects as a function of the masker azimuth in degrees. The position of the probe speaker was fixed at $-90°$. The diamond indicates the average threshold for detecting probe modulation without the masker present. The circles represent thresholds obtained in the presence of the unmodulated masker and the triangles represent conditions in which the masker was modulated. Error bars represent the standard error of the mean threshold.

4. Subjects

Twelve new subjects were recruited for experiment 2. Their pure-tone thresholds were within 10 dB of ANSI (1989) standard audiometric thresholds. Six subjects participated in the first set of conditions and the other six in the second set. Subjects received a minimum of 16 h of practice before data collection began, including 4 h of practice in the sound-field conditions.

B. Results and discussion

1. Effect of probe-to-masker source separation

Figure 3 shows mean AM detection thresholds for the six subjects as a function of the lateral position of the masker speaker. The probe was always presented through the speaker positioned at $-90°$ directly facing the subject's left ear. The diamond indicates the average threshold for detecting 10-Hz probe modulation without the masker present. The circles represent thresholds obtained in the presence of the unmodulated masker. Masker position had little to no effect on performance with the mean unmodulated-masker thresholds varying by less than 1 dB as a function of masker location.

The triangles represent thresholds obtained in the presence of the modulated masker. With the modulated masker there was greater change in threshold as a function of masker lateral position than in the unmodulated-masker conditions. The mean threshold dropped by roughly 3 dB when the masker position was moved from the probe position of $-90°$ to $-45°$ with no further improvement in performance obtained with a larger spatial separation between sources. If

spatial separation promotes perceptual segregation of the probe and masker to reduce MDI, the greatest effect would be expected when changing from one to two source locations. A 3-dB change in threshold as a function of masker lateral position suggests that spatial separation of the probe and masker does not have a strong effect on MDI.

A repeated-measures ANOVA was performed on factors masker modulation index and masker position. There was a significant effect of modulation index ($F_{1,5}=53.9$, $p=0.001$), masker position ($F_{3,15}=3.8$, $p=0.033$) and a significant interaction between the two factors ($F_{3,15}=5.7$, $p=0.008$). *Post-hoc* pairwise comparisons were performed among the four positions at each level of modulation index. For the unmodulated masker, no comparisons were significant. With the modulated masker, the significant comparisons were between masker position at $-90°$ and the three other positions.

Sound-field thresholds show less variation across the modulated-masker conditions than obtained in the IID conditions of experiment 1. Under headphones, performance was poorest with diotic presentation of the probe and masker and best when the probe and masker were presented to separate ears. Due to the frequency dependence of the directional response of the ear as a function of azimuth (Shaw, 1974), the IIDs for the probe and masker were never identical in any of the sound-field conditions. That is, the sound-field conditions would not replicate the monotic or diotic reference for MDI used in the headphone study. Also, sound-field conditions cannot duplicate the headphone condition with the probe and masker only stimulating opposite ears. Thus the sound-field conditions cannot replicate the IID disparities of the conditions which gave poorest and best performance in experiment 1 with the sound-field IIDs presumably falling between the extremes of the headphone conditions. Actual IIDs were not measured at the subjects' eardrum. When allowing for the possibility of small head movements by the subject, Shaw's (1974) data suggest that a 45° separation between the probe and masker position in the sound field could introduce a sufficient disparity in the probe and masker IIDs to reduce the amount of interference relative to that obtained with a single source for the probe and masker. As in the IID conditions of experiment 1, a disparity between the IIDs of the probe and masker results in a monotic level difference between the probe and masker. This level difference, and not perceptual segregation due to source location, may have been the basis for the small effect of spatial separation on MDI.

2. Effect of masker AM rate

In the first set of sound-field conditions, the probe and masker modulation rates were the same. In most multisource listening environments, sources would be characterized by different patterns of envelope modulation. Results from previous studies in which stimuli were presented over headphones show that MDI can be affected by a difference in the modulation rate of the probe and masker (Yost *et al.*, 1989; Moore *et al.*, 1991; Bacon and Opie, 1994; Bacon *et al.*, 1995). The amount of interference tends to gradually decline with increasing separation between the two modulation rates.
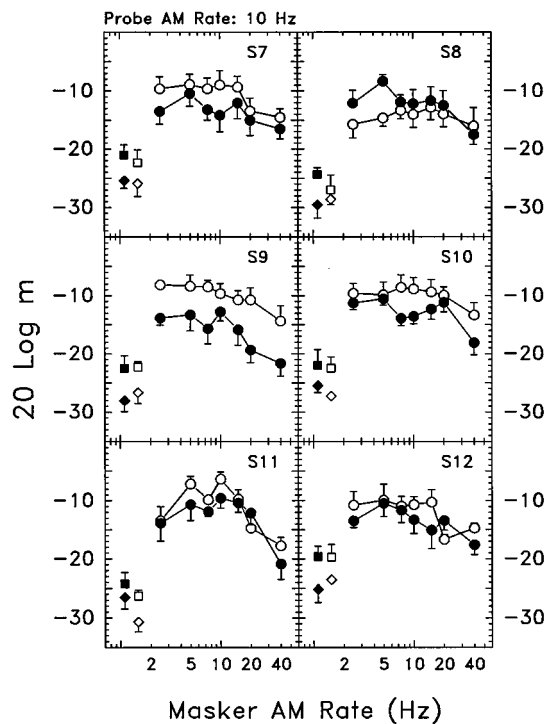
FIG. 4. Thresholds for detecting 10-Hz probe modulation as a function of the masker AM rate. The open and filled symbols represent results obtained under headphones and in the sound field, respectively. In the sound-field conditions, the position of the probe speaker was fixed at 0° azimuth and the masker at 90°. The diamonds and squares on the left side of each panel represent the subject's thresholds for detecting probe modulation without the masker present and in the presence of the unmodulated masker, respectively. Error bars represent one standard deviation from the mean threshold.

Bacon and Opie (1994) compared performance in monotic to dichotic conditions. For their practiced subjects, the tuning in the modulation domain was on average even broader when the probe and masker were presented to separate ears. In the individual data, the patterns for the change in threshold as a function of masker modulation rate showed differences between the monotic and dichotic conditions. The second set of conditions in experiment 2 compared the effect of masker AM rate in sound-field conditions to results obtained with diotic stimulus presentation over headphones.

Figure 4 shows thresholds for detecting 10-Hz probe modulation as a function of the masker AM rate. The open symbols represent results obtained under headphones and the filled symbols indicate thresholds obtained in the sound field. In the sound-field conditions, the position of the probe speaker was fixed at 0° azimuth and the masker at 90°. The diamonds at the far left of each panel indicate the subject's thresholds for detecting probe modulation without the masker present and the squares represent conditions in which the masker was present but not modulated. In both the probe-alone and unmodulated-masker conditions, the sound-field and headphone thresholds were generally within 2 dB of each other with the largest difference of roughly 4 dB obtained in the probe-alone thresholds of subject S11.

The pattern of the change in threshold as a function of masker AM rate varied across subjects. Looking first at the results from the headphone conditions (open circles), only the thresholds from subject S11 showed a bandpass charac-

teristic with the greatest amount of interference when the probe and masker AM rates were the same. For the other subjects, the functions were low pass or close to flat with little change in performance as a function of the masker modulation rate. Three previous studies have measured thresholds for detecting 10-Hz probe modulation in the presence of maskers modulated at various rates (Yost *et al.*, 1989; Bacon and Opie, 1994; Bacon *et al.*, 1995). Across studies, the degree of high-pass tuning in the functions varies. All but one subject in the present study, and also the results from Yost *et al.* (1989), showed little or no improvement in performance as the masker AM rate was lowered from the probe modulation rate of 10 Hz. Bacon and Opie (1994) and Bacon *et al.* (1995) reported an average 4- to 5-dB drop in threshold with decreasing masker modulation rate from 10 Hz as long as a masker carrier frequency was higher than the probe carrier frequency. Though procedural differences (e.g., number of masker carriers and carrier frequencies) may in part be responsible for the discrepancy among studies, the discrepancy may also relate to the large individual differences common in measures of MDI.

The primary goal of the present experiment was to compare performance in the sound field to headphone conditions. The filled circles in Fig. 4 show the sound-field thresholds as a function of the masker AM rate. The sound-field thresholds tend to either overlap or fall below the results from the headphone conditions. The most notable exception was for subject S8 with masker AM rates of 2.5 and 5 Hz. When referenced to the unmodulated-masker thresholds, the amount of MDI was therefore often smaller in the sound-field conditions. The sound-field mode of presentation was dichotic with the probe speaker at 0° azimuth and the masker at 90°, whereas the headphone mode of presentation was diotic. Results from the IID conditions of experiment 1 suggest that this contrast between dichotic and diotic modes of presentation may have been the basis for the smaller amounts of MDI obtained in the sound-field conditions.

For subjects S7, S8, S10, and S12 in the sound-field conditions, MDI was not greatest when the probe and masker AM rates were both 10 Hz, with thresholds slightly poorer when the masker AM rate was below 10 Hz. However, to a first approximation, the extent of tuning in the modulation domain was similar for results obtained under headphones and in the sound field. That is, spatial separation did not appear to substantially alter the effect of a disparity between the probe and masker modulation rates on MDI.

III. SUMMARY

Results from the present study show that binaural stimulus parameters can affect MDI. When stimuli were presented over headphones, the introduction of interaural level, but not time, differences reduced the amount of interference. In the conditions studied, a disparity between the IID of the probe and masker resulted in a monotic level difference between the probe and masker at one or both ears. The similarity in thresholds obtained in conditions with 6-dB IIDs and monotic conditions with a 6-dB level difference suggests that monotic level differences may have been a contributing factor to the reduction in MDI in the dichotic conditions. Re-

sults obtained with presentation of the probe and masker through separate speakers in a sound-deadened room indicate that spatial separation between the sources also led to a reduction in MDI. The improvement in performance with spatial separation of the sources was consistent with the range of threshold variation obtained in the IID conditions when testing under headphones. As in the headphone conditions, monotic level differences between the probe and masker in the sound-field conditions may have contributed to the effect of spatial separation on MDI. Comparison of sound-field and headphone results indicates that spatial separation does not appreciably change the extent of tuning in the modulation domain. Overall, results indicate that segregation of the probe and masker based on binaural cues does not lead to a substantial reduction in the amount of interference.

ACKNOWLEDGMENTS

[1]Though subjects were informed of the expected lateral position of the probe and masker in each condition and their report was generally consistent with expectation, no formal effort was made to determine the perceived probe and masker lateral positions across conditions.

ANSI (**1989**). ANSI S3.6-1989, ''Specifications for audiometers'' (American National Standards Institute, New York).

Bacon, S. P., and Konrad, D. L. (**1993**). ''Modulation detection interference under conditions favoring within- or across-channel processing,'' J. Acoust. Soc. Am. **93,** 1012–1022.

Bacon, S. P., and Moore, B. C. J. (**1993**). ''Modulation detection interference: Some spectral effects,'' J. Acoust. Soc. Am. **93,** 3442–3453.

Bacon, S. P., and Opie, J. M. (**1994**). ''Monotic and dichotic modulation detection interference in practiced and unpracticed subjects,'' J. Acoust. Soc. Am. **95,** 2637–2641.

Bacon, S. P., Moore, B. C. J., Shailer, M. J., and Jorasz, U. (**1995**). ''Effects of combining maskers in modulation detection interference,'' J. Acoust. Soc. Am. **97,** 1847–1853.

Bregman, A. S. (**1990**). *Auditory Scene Analysis* (MIT, Cambridge, MA).

Buell, T. N., and Hafter, E. R. (**1991**). ''Combination of binaural information across frequency bands,'' J. Acoust. Soc. Am. **90,** 1894–1900.

Cramer, E. M., and Huggins, W. H. (**1958**). ''Creation of pitch through binaural interaction,'' J. Acoust. Soc. Am. **30,** 413–417.

Dye, R. H. (**1990**). ''The combination of interaural information across frequencies: Lateralization on the basis of interaural delay,'' J. Acoust. Soc. Am. **88,** 2159–2170.

Dye, R. H. (**1997**). ''The relative contributions of targets and distractors in judgements of laterality based on interaural differences of level,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum, Mahwah, NJ), pp. 151–168.

Dye, R. H., Yost, W. A., Stellmack, M. A., and Sheft, S. (**1994**). ''Stimulus classification procedure for assessing the extent to which binaural processing is spectrally analytic or synthetic,'' J. Acoust. Soc. Am. **96,** 2720–2730.

Dye, R. H., Stellmack, M. A., Grange, A. N., and Yost, W. A. (**1996**). ''The effect of distractor frequency on judgements of target laterality based on interaural delays,'' J. Acoust. Soc. Am. **99,** 1096–1107.

Hall, J. W., and Grose, J. H. (**1991**). ''Some effects of auditory grouping factors on modulation detection interference (MDI),'' J. Acoust. Soc. Am. **90,** 3028–3035.

Kubovy, M., and Daniel, J. E. (**1983**). ''Pitch segregation by interaural phase, by momentary amplitude disparity, and by monaural phase,'' J. Audio Eng. Soc. **31,** 630–634.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49,** 467–477.

Mendoza, L., Hall, J. W., and Grose, J. H. (**1995**) ''Within- and across-channel processes in modulation detection interference,'' J. Acoust. Soc. Am. **97,** 3072–3079.

Moore, B. C. J., and Shailer, M. J. (**1992**). ''Modulation discrimination interference and auditory grouping,'' Philos. Trans. R. Soc. London, Ser. B **336,** 339–346.

Moore, B. C. J., and Shailer, M. J. (**1994**). ''Effects of harmonicity, modulator phase, and number of masker components on modulation discrimination interference,'' J. Acoust. Soc. Am. **95,** 3555–3560.

Moore, B. C. J., Glasberg, B. R., Gaunt, T., and Child, T. (**1991**). ''Across-channel masking of changes in modulation depth for amplitude- and frequency-modulated signals,'' Q. J. Exp. Psych. **43A,** 327–347.

Shailer, M. J., and Moore, B. C. J. (**1993**). ''Effects of modulation rate and rate of envelope change on modulation discrimination interference,'' J. Acoust. Soc. Am. **94,** 3138–3143.

Shaw, E. A. G. (**1974**). ''Transformation of sound-pressure level from the free field to the eardrum in the horizontal plane,'' J. Acoust. Soc. Am. **56,** 1848–1861.

Sheft, S. (**1997**). ''Effects of carrier asynchrony on modulation detection interference,'' J. Acoust. Soc. Am. (submitted).

Stellmack, M. A., and Dye, R. H. (**1993**). ''The combination of interaural information across frequencies: The effects of number and spacing of components, onset asynchrony, and harmonicity,'' J. Acoust. Soc. Am. **93,** 2933–2947.

Stellmack, M. A., and Lutfi, R. A. (**1996**). ''Observer weighting of concurrent binaural information,'' J. Acoust. Soc. Am. **99,** 579–587.

Yost, W. A. (**1981**). ''Lateral position of sinusoids presented with interaural intensive and temporal differences,'' J. Acoust. Soc. Am. **70,** 397–409.

Yost, W. A. (**1991a**). ''Thresholds for segregating a narrow-band from a broadband noise based on interaural phase and level differences,'' J. Acoust. Soc. Am. **89,** 838–844.

Yost, W. A. (**1991b**). ''Auditory image perception and analysis: The basis for hearing,'' Hearing Res. **56,** 8–18.

Yost, W. A., and Sheft, S. (**1989**). ''Across-critical-band processing of amplitude-modulated tones,'' J. Acoust. Soc. Am. **85,** 848–857.

Yost, W. A., and Sheft, S. (**1990**). ''A comparison among three measures of cross-spectral processing of amplitude modulation with tonal signals,'' J. Acoust. Soc. Am. **87,** 897–900.

Yost, W. A., and Sheft, S. (**1994**). ''Modulation detection interference: Across-frequency processing and auditory grouping,'' Hearing Res. **79,** 48–58.

Yost, W. A., Sheft, S., and Opie, J. (**1989**). ''Modulation interference in detection and discrimination of amplitude modulation,'' J. Acoust. Soc. Am. **86,** 2138–2147.

The effects of frequency region and bandwidth on the temporal modulation transfer function[a]

Elizabeth A. Strickland[b] and Neal F. Viemeister
Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455

Temporal resolution was examined as a function of frequency region and listening region. The first experiment demonstrated that amplitude- and frequency-modulated tones are not appropriate stimuli to study temporal resolution as a function of frequency region, due to the availability of other cues in addition to temporal ones. In the other experiments, thresholds for detection of sinusoidal amplitude modulation of a noise band were measured as a function of frequency region, bandwidth, and level of surrounding notched noise masker. Temporal modulation transfer functions (TMTFs) measured in low- and high-frequency regions did not differ in sensitivity or in cutoff frequency, suggesting that initial ''critical band'' filtering did not affect temporal resolution. When the upper cutoff frequency of the noise was held constant, TMTF sensitivity increased with noise bandwidth, while the cutoff frequency of the TMTF did not show measurable change. These results are consistent with the predictions of an envelope detector model if peripheral filtering in the lower-frequency range is assumed to be approximately twice as wide as that estimated by measuring thresholds for a tone in notched noise. Restricting the listening region with notched noise increased thresholds for low modulation frequencies but not for high. This is consistent with other data showing that upward spread of excitation may increase the effective modulation depth, but only for low modulation frequencies. © *1997 Acoustical Society of America.* [S0001-4966(97)04109-X]

PACS numbers: 43.66.Mk, 43.66.Ba [RHD]

INTRODUCTION

The peripheral auditory system functions as a bank of bandpass filters, with bandwidths that increase nearly proportionally with center frequency. Because the time constant of a linear filter is inversely proportional to its bandwidth, the ability to follow temporal changes in a sound would be expected to increase with peripheral filter width, and thus with frequency, if peripheral filtering is the limiting factor. Physiological data are consistent with this prediction up to a point. Measurements of synchrony to temporal envelopes in cat auditory nerve fibers show an increase in temporal resolution with fiber characteristic frequency (CF) up to approximately 10 kHz (Joris and Yin, 1992), and are constant at higher CFs. This type of relationship suggests that there may be peripheral and central limitations to temporal resolution, an idea captured in a model which will be introduced later. In this case, peripheral filtering would only be the limiting factor for temporal resolution in frequency regions in which it was more severe than the effective central limitation. The presence of both a peripheral and a central limitation may explain why a relationship between frequency region and temporal resolution has not been clearly demonstrated psychophysically (e.g., Formby and Muir, 1988).

A characterization of the temporal resolution of intensity changes by the auditory system may be obtained by measuring a temporal modulation transfer function (TMTF). The

TMTF can be determined by measuring thresholds for detection of sinusoidal amplitude modulation (SAM) of a carrier as a function of modulation frequency (Viemeister, 1977), and resembles a lowpass filter function. This measure is particularly useful compared to other measures such as gap detection, because it allows the effects of intensity resolution and temporal resolution to be observed separately. Intensity resolution determines threshold at very low modulation frequencies, where temporal resolution is not a factor. The absolute value of this low-frequency plateau will be referred to as sensitivity (S). Temporal resolution increases with the 3-dB cutoff frequency of the function, F_c, when it is fit by a lowpass filter function.

The goal of this study is to measure TMTFs in different frequency regions, to further examine the relationship between temporal resolution and peripheral filtering.

I. EXPERIMENT I. DISCRIMINATION OF AM AND QFM TONES

It might seem that the easiest way to control frequency region would be to use tonal carriers. In this case, the waveforms are deterministic, avoiding the possible effects of inherent fluctuations of noise carriers, which might interfere with perception of modulation. The amplitude spectrum of a SAM tone contains a component at the carrier frequency and sidebands at the carrier frequency plus and minus the modulation frequency, as shown in Eq. (1):

$$\text{AM}(t) = \cos{(2\pi f_c t)} - \frac{m}{2}\left[\cos{(2\pi(f_c - f_m)t)}\right.$$
$$\left. + \cos{(2\pi(f_c + f_m)t)}\right], \tag{1}$$

where m is the modulation index, f_c is the carrier frequency, and f_m is the modulation frequency. To use a SAM tone to measure amplitude modulation detection, the appropriate comparison would have the same amplitude spectrum but a flat envelope. The closest approximation to this is a quasi-frequency modulated (QFM) tone, as defined in Eq. (2). Spectrally, it differs from the AM tone only in that the sidebands are in 90° phase relative to the carrier. It is called QFM because the instantaneous frequency of the tone changes over time:

$$\text{FM}(t) = \cos\left(2\pi f_c t\right) + \frac{m}{2}\left[\cos\left(2\pi(f_c - f_m)t + \frac{\pi}{2}\right)\right.$$
$$\left. + \cos\left(2\pi(f_c + f_m)t + \frac{\pi}{2}\right)\right]. \tag{2}$$

In the case of the QFM tone, m is a scaling factor for the sidebands, but does not correspond to modulation index in the way that it does for the AM tone.

Previous research has suggested that several cues may be used for discriminating AM from QFM (Buunen, 1976). These include

(1) the temporal envelope, which has a greater modulation depth for AM than QFM, and which has a frequency of f_m for AM and $2f_m$ for QFM;
(2) frequency sweeps in the QFM signal; and
(3) a difference in the amplitude of the lower sideband in the internal spectrum, caused by the cubic difference tone from the carrier and upper sideband adding to the lower sideband in a different phase for AM than for QFM.

To obtain a valid TMTF, one that reflects temporal processing, only the first cue should be available. In order to determine whether the other cues are available, in Experiment I discrimination was measured for AM and QFM tones as a function of modulation frequency, and also for QFM versus a pure tone at the carrier frequency. Assume that for a given modulation frequency, the discrimination threshold [$20 \log(m)$] for AM versus QFM is well below that for QFM versus tone. Then it may be assumed that the AM versus QFM discrimination is being made on the basis of envelope fluctuations and not on the basis of any other cues in QFM, because QFM is essentially indiscriminable from a pure tone. Now assume that the discrimination threshold for QFM versus pure tone is close to or lower than that for AM versus QFM. In this case, the basis for discrimination between AM and QFM is no longer clear, because any of the cues mentioned above may be available.

A. Procedure

A 4-kHz pure tone and AM and QFM tones with a 4-kHz carrier were digitally generated, and played at a rate of 44.1 kHz. They were lowpass filtered (Kemo VBF/25) at 10 kHz at an attenuation rate of 135 dB/oct. The AM and QFM waveforms were scaled by a factor of $(1 + m^2/2)^{-1/2}$ to equalize the overall power to that of the pure tone. The level was set at 70 dB SPL. The signal duration was 500 ms, including 10-ms $\cos^2$ ramps. Subjects were tested individually in a sound-attenuating booth through TDH49 earphones
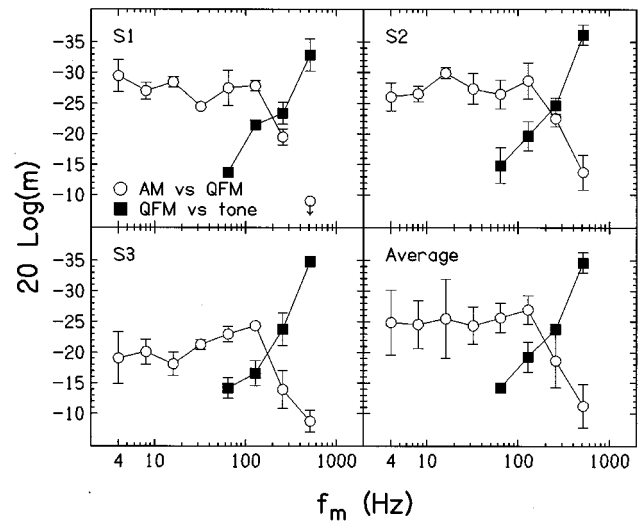
FIG. 1. Thresholds as a function of f_m for AM versus QFM and QFM versus a tone for three subjects and their average. Error bars are the standard deviation. The symbol with the downward arrow (upper left panel) indicates that no threshold could be measured due to termination of the track.

in GS001 cushions. An adaptive 3IFC task with a two-down, one-up stepping rule was used to measure the level of the sidebands necessary for 70.7% correct discrimination (Levitt, 1971). The step size was 4 dB for the first four reversals, then decreased to 2 dB. Of 12 total reversals, the last 8 were averaged for threshold. If a step would have caused modulation to exceed 100%, the track was terminated. Runs with standard deviations larger than 3.5 dB were discarded. At least three runs were averaged for each threshold estimate.

B. Subjects

Three females served as subjects. They had pure tone thresholds within 10 dB of laboratory norms at octave frequencies from 250–8000 Hz in the ear tested.

C. Results

Data for the three subjects are plotted in terms of m in dB in Fig. 1. Open circles show data for AM versus QFM discrimination, and filled squares for QFM versus tone discrimination. Error bars are standard deviations. The AM versus QFM function is remarkably flat up to $f_m = 128$ Hz, compared to TMTFs for noise carriers (e.g., Viemeister, 1979). Thresholds for discriminating QFM from a tone decrease as f_m increases. At and above 128 Hz, where the AM versus QFM function starts to roll off, discrimination thresholds for QFM versus a tone approach and then become lower than AM versus QFM discrimination thresholds (higher on the figure). The flatness of the function and the frequency at which AM versus QFM performance starts to decrease is similar to data from Fassel and Kohlrausch (1996) for 3- and 5-kHz carriers. From these data it is impossible to determine the relative contributions of the three cues mentioned above for f_m's just below and just above 128 Hz. Even though whatever cue is being used in the QFM versus tone discrimination is below threshold of 70.7% correct, it might still be
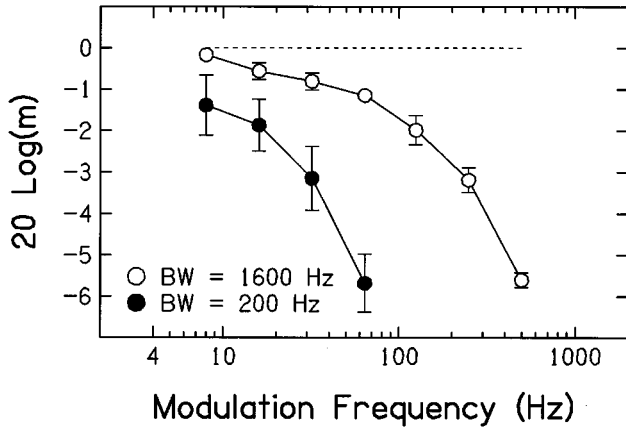
FIG. 2. Modulation depth for a broadband noise that is 100% SAM (indicated by the dashed line) then filtered to a bandwidth of 1600 and 200 Hz.

used in combination with another cue in discriminating AM from QFM. This could explain the unusual flatness of the function up to this point. Therefore the AM versus QFM function cannot be simply interpreted as a TMTF for f_m at and above 128 Hz. The point of this section is simply to demonstrate that this set of deterministic stimuli cannot be used to measure a TMTF that validly reflects temporal resolution in the frequency regions of interest in this study.

II. EXPERIMENT II: DISCRIMINATION OF AM AND QFM NOISE

A. Rationale

Previous studies have used spectrally limited noise bands to measure the TMTF for a specific frequency region. In most research using this technique, the bandwidth and frequency region of the noise bands have been increased together (Viemeister, 1979; Formby and Muir, 1988; van Zanten, 1980; Rodenburg, 1977). Therefore the effects of bandwidth were confounded with those of frequency region. Eddins (1993) used sharp filtering to vary bandwidth independently of frequency region. He concluded that the TMTF did not vary with frequency region, but that both sensitivity and temporal resolution increased with bandwidth. In his study, as in most of the previous studies, filtering was done *after* modulation, so that there would be no spectral cues to modulation. However, such filtering could decrease the modulation depth of the stimuli. Figure 2 shows data from a simulation in which noise was 100% SAM and then rectangularly filtered to a bandwidth of 200 or 1600 Hz. The envelope was obtained using the Hilbert transform, and the modulation depth was measured in units of $20 \log(m)$, where m is the ratio of the amplitude of the envelope component at the modulation frequency to the envelope DC. The dotted line indicates the modulation depth of the noise before filtering, which is $20 \log(m) = 0$, or 100% modulation. Notice that the actual modulation depth of the filtered noise depends on both the modulation frequency and the bandwidth, and modulation depth is reduced by 3 dB at a modulation frequency that is approximately $\frac{1}{8}$ of the filter bandwidth. Although the modulation depth could be corrected for the effects of filtering, there is a more direct approach, one that extends the idea of using AM and QFM tones, but that eliminates the problems demonstrated in experiment I.

According to the Rice model (Rice, 1944), a narrowband noise may be approximated by a sum of sinusoids, as in Eq. (3):

$$N(t) = \sum_{c=l}^{h} A_c \cos(2\pi f_c t + \phi_c), \tag{3}$$

where A_c is an amplitude randomly chosen from a Rayleigh distribution, ϕ_c is a random phase rectangularly distributed between 0 and 2π radians, f_c is the frequency of each component, and f_l and f_h are the lower and upper spectral edges of the noise. In order to amplitude modulate the noise, each sinusoid is now treated as a carrier, and sidebands in $-\cos$ phase with the carrier are added at plus and minus the modulation frequency:

$$AM(t) = \sum_{c=l+m}^{h-m} A_c \left[\cos(2\pi f_c t + \phi_c) - \frac{m}{2}(\cos(2\pi(f_c - f_m)t + \phi_c) + \cos(2\pi(f_c + f_m)t + \phi_c)) \right], \tag{4}$$

where m is the modulation index (0 to 1), f_m is the modulation frequency, $f_{l+m} = f_l + f_m$, $f_{h-m} = f_h - f_m$, and the indices c, l, m, and h are harmonics of 1 Hz. The lower and upper spectral edges of the carrier band are increased and decreased, respectively, by f_m, so that when the sidebands are added, the total bandwidth of the signal stays constant at $f_h - f_l$, regardless of the modulation frequency. Thus the carrier bandwidth decreases with f_m, and the sidebands increase with f_m.

To create a comparison noise that is not amplitude modulated, two options are considered. The same sidebands may be shifted to 90° phase with respect to their carriers and added to create what will be called QFM noise:

$$QFM(t) = \sum_{c=l+m}^{h-m} A_c \left[\cos(2\pi f_c t + \phi_c) \right.$$
$$- \frac{m}{2}\left(\cos\left(2\pi(f_c - f_m)t + \phi_c + \frac{\pi}{2}\right)\right.$$
$$\left.\left. + \cos\left(2\pi(f_c + f_m)t + \phi_c + \frac{\pi}{2}\right)\right)\right]. \tag{5}$$

The sidebands may also be added in random phase, which will be called random noise:

$$Random(t) = \sum_{c=l+m}^{h-m} A_c \left[\cos(2\pi f_c t + \phi_c) \right.$$
$$- \frac{m}{2}(\cos(2\pi(f_c - f_m)t + \theta_c)$$
$$\left. + \cos(2\pi(f_c + f_m)t + K_c)) \right], \tag{6}$$

where Θ_c and K_c are independent random phases. Examples of the amplitude spectra and waveforms for each type of
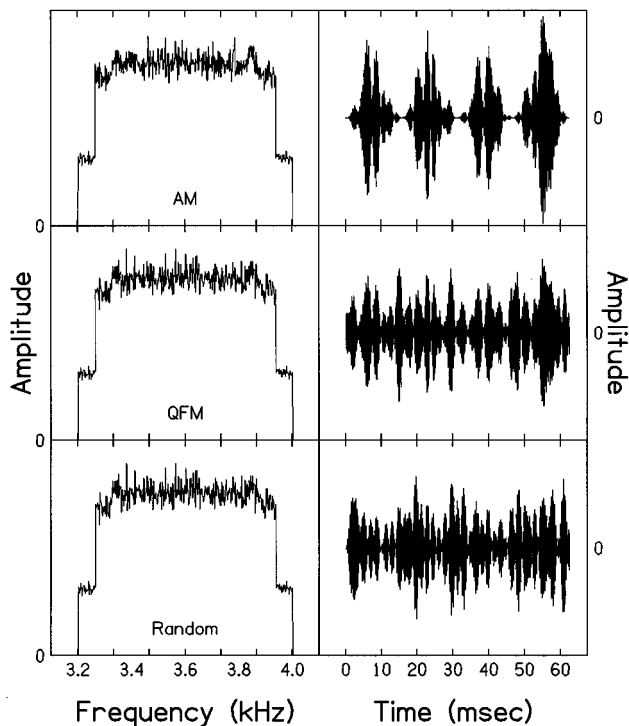
FIG. 3. Amplitude spectra and waveforms for AM, QFM, and random noise (100 averages). The bandwidth is 800 Hz, f_h is 4 kHz, and f_m is 64 Hz.

noise are shown in Fig. 3, for $f_l = 3.2$ kHz, $f_h = 4$ kHz, $f_m = 64$ Hz, and $m = 1$.

Generating noise bands in this manner eliminates the problems of filtering after modulation. In Fig. 2, it was shown that the modulation depth of a narrowband noise filtered after modulation depends on the modulation frequency and the bandwidth of the signal. Noise bands generated using Eq. (4) have modulation depths that are independent of bandwidth. Using noise bands should also eliminate some of the problems found with tones, mentioned in the previous section. In particular, instantaneous frequency changes should be less detectable for a noise band, due to the fact that its pitch is less distinct than that of a tone, and the frequency components are randomized on each trial. Pilot data for subject 1 indicated that QFM noise was indiscriminable from random noise except at very high modulation frequencies, and at these frequencies the thresholds were higher than those subsequently found for discrimination of AM and QFM noise, suggesting that instantaneous frequency cues were not the basis for discrimination in the experiments.

Although previous studies have controlled the frequency region of the signal, they have not limited the listening region which is available. It is possible that listeners could attend to high-frequency filters well above the signal frequencies, which pass the signals through their low-frequency tails, regardless of the signal frequency. In this case, temporal resolution would be determined by the central limitation, and there would be no effect of frequency region of the signal. In addition, at frequency regions above the signal frequency, nonlinear upward spread of excitation might effectively magnify the modulation depth. This second possibility

has been examined in previous research by measuring masking produced by a SAM tonal masker for probes at and above the masker frequency (Zwicker, 1976; Arndt, 1993, see also Nelson, 1992), as a function of the temporal position of the probe within peaks or valleys of the masker envelope. This is called the temporal probe technique (Viemeister, 1977). The difference between peak and valley thresholds may be used to calculate the effective modulation depth of the masker in different frequency regions. The data of Zwicker and of Arndt indicate that for low modulation frequencies, the effective modulation depth is greater in the frequency region above the masker frequency than at the masker frequency. The effective modulation depth decreases with masker modulation frequency, and is almost zero at a masker modulation frequency of 64 Hz. This suggests that, at least at low modulation frequencies, listeners might be able to use upward spread of excitation to increase the effective modulation depth. Therefore, another aim of the present study was to examine the effects of truly limiting the listening region, by surrounding the signal region with masking noise.

B. Condition 1. Effects of masking noise spectrum and level on the TMTF

An experiment was done to determine the effect of masking noise spectrum and level on the sensitivity (S) of the TMTF. If upward spread of excitation is used in detecting SAM, then masking noise above the test stimuli would be expected to increase modulation thresholds.

1. Stimuli

The test stimuli (noise bands) and masking noise were digitally generated by a NeXT computer, according to the formulas presented above. The spacing of frequency components was 1 Hz. In generating the test stimuli, three separate 1.5-s files were generated; one containing the carrier band of noise, one with the sidebands in AM phase, and a third with the sidebands in random phase. Random samples were chosen from the carrier file for each interval, with samples chosen from the same point in the sideband files to preserve the phase relationship. For the modulated noise, the samples were chosen so that the envelope started in $-$cos phase.

The upper cutoff frequency of the test stimuli was 2 kHz, the bandwidth ($f_h - f_l$) was 800 Hz, and the modulation frequency was 8 Hz. The overall level was set to be the same as that of a flat-spectrum noise of the same bandwidth at a spectrum level of 50 dB, i.e., 79 dB SPL.

Low- and high-frequency masking bands were generated. The low-frequency band extended from 10–1150 Hz (50 Hz below f_l), and the high-frequency band extended from 2050 Hz (50 Hz above f_h) to 10 kHz. The masking noise was set at spectrum levels from -20 dB (at which it was barely audible) to 40 dB.

The stimulus duration was 500 ms, including 10-ms $\cos^2$ ramps. The masking noise was gated with the test stimuli. The masking noise was played at a rate of 44.1 kHz. The test stimuli were played at an effective rate of 22.05 kHz; that is, there were 44 100 samples/s, but even-
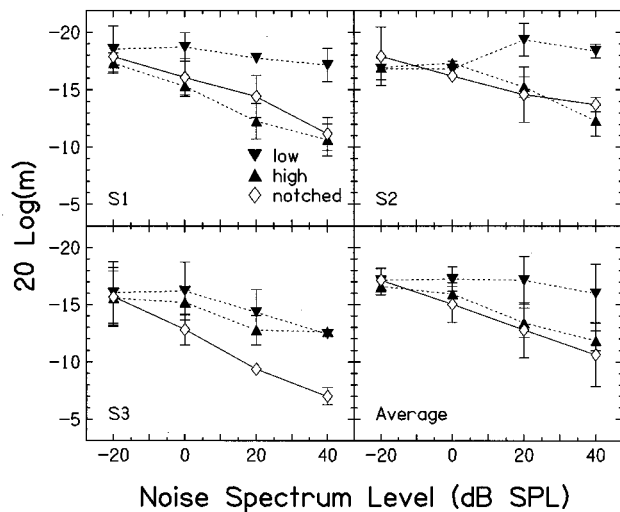
FIG. 4. Thresholds as a function of masking noise level for three subjects and their average. The test stimulus is 800 Hz wide, f_h is 2 kHz, and f_m is 8 Hz. The spectrum level of the test stimulus is 50 dB SPL. Masking noise was positioned below (low), above (high), or both below and above (notched) the test stimulus.

numbered samples were simply duplicates of the preceding odd-numbered samples. Lowpass filtering, tracking, and subjects were the same as in experiment I.

2. Results

In Fig. 4, the data for three subjects and their average are shown as a function of the spectrum level of the masking noise. Downward-pointing triangles show data for the low-frequency masking band, and upward-pointing triangles are for the high-frequency masking band. Data for the two bands combined are shown by the diamonds.

For all three subjects, modulation detection thresholds increased nearly linearly (downward on the figure) with masking noise level when the two bands were present. A linear fit to the average two-band data had a slope of 0.11 ($r=1.0$). For S1 and S2, the low-frequency band had little or no effect on thresholds, while the thresholds with the high-frequency band were nearly the same as those when both bands were present. S3 showed small, nearly equal increases in threshold with low- or high-frequency noise. On average, however, the high-frequency band appeared to do almost all of the masking. This suggests that subjects may have been able to listen above the signal frequency when no masking noise was present, and that this made modulation more detectable.

C. Simulations to determine signal parameters

Simulations were done in order to choose conditions which would be expected to show the effects of peripheral filtering. In order to quantify the expected effects of noise bandwidth and frequency on the form of the TMTF, a model which will be called an ''envelope detector'' was used. This model, consisting of one or more bandpass filters, a half-wave rectifier, and a lowpass filter, has been used in previous research based on linear systems analysis (Rodenburg, 1977;

Viemeister, 1977, 1979), and is a modification of a type of model that has been commonly used in psychophysics (e.g., Jeffress, 1967). A fourth-order gammatone filter was used for the initial bandpass filtering (Patterson *et al.*, 1987). This filter approximates auditory filters which have been derived from masking data in which the level of a notched-noise masker is fixed, and the threshold for a tone is measured as a function of the position and width of the notch in the noise (e.g., Patterson and Moore, 1986). The equivalent rectangular bandwidth (ERB) of the gammatone filter was set to approximate that of an auditory filter measured by the notched-noise method, using the ERB formula proposed by Glasberg and Moore (1990). The optimally placed bandpass filter was chosen based on modulation depth at the output of the lowpass filter. The lowpass filter was a first-order Butterworth with a cutoff frequency of 90 Hz. This cutoff frequency was set by the 3-dB cutoff frequency of a lowpass filter function fit to an average TMTF for a continuous broadband noise carrier, measured in a previous study (Strickland and Viemeister, 1996).

The first question was whether to use QFM or random noise as the comparison. The QFM noise was found to be preferable for theoretical reasons.[1] In the simulations, AM and QFM noise bands were added to masking noise with a notch extending from 50 Hz above to 50 Hz below the spectral edges of the stimuli, and a spectrum level 10 dB below the spectrum level of the noise bands, as in the highest masking noise level in condition 1. The decision statistic was the modulation depth at the output of the envelope detector in units of $20 \log(m)$, where m is the ratio of the amplitude of the envelope component at the modulation frequency to the envelope component at 0 (the dc).

Several factors determined the final choice of stimulus parameters. In order to examine the effects of noise band frequency on the form of the TMTF, the goal was to choose two frequency regions which were as far apart as possible to maximize the effects of peripheral filtering. The headphones used in the experiments were TDH49's, which have a frequency response curve that begins to roll off gradually above 5 kHz (as measured on a KEMAR). Because we wanted to examine the effects of notched noise, the upper cutoff frequency (f_h) of the higher-frequency noise band was set at 4 kHz, to ensure that sufficient levels of notched noise would be passed above it to prevent off-frequency listening. For the lower-frequency noise band, f_h had to be set so that the entire bandwidth was within the audible range, with a bandwidth wide enough to allow high modulation frequencies to be measured which would show the effects of peripheral filtering. Pilot data also indicated that the lower cutoff frequency of the lower-frequency noise band should be above the highest modulation frequency measured, because it was necessary to place a masker at low frequencies to mask an audible distortion product at the difference tone (which is the modulation frequency). This distortion product was produced in the ear, and was not present in the acoustic waveform. After trial and error, it was determined that a bandwidth of 800 Hz and an f_h of 1.2 kHz for the lower-frequency test stimulus should show the effects of peripheral filtering, if they were present. According to the simulations, if peripheral

filtering limits temporal resolution, the TMTF for the lower-frequency signal should roll off faster than that for the higher-frequency signal, diverging by as much as 15 dB at a modulation frequency of 256 Hz. Predictions of a slightly different implementation of the model are shown in Fig. 9, which will be discussed later.

D. Conditions 2 and 3

1. Method

Two of the subjects from the previous experiments and two new subjects participated. Stimulus generation and procedure were the same as in condition 1 except as noted here. The QFM noise was presented in the nonsignal interval. The overall level of the test stimuli was the same as that of a flat-spectrum noise of the same bandwidth, at a spectrum level of 50 dB. The masker was a notched noise consisting of both the low- and high-frequency bands from condition 1. An additional band of masking noise extending from 100–300 Hz was generated in the same manner as the notched noise, with a spectrum level of 50 dB. This noise was presented when the spectrum level of the notched noise was -20 dB, to mask a difference tone that was audible at high modulation frequencies, discussed above. It was gated with the test stimuli.

The initial step size was 2 dB and was halved after four reversals. Of 16 total reversals, the final 12 were averaged to estimate threshold. If a step would have caused modulation to exceed 100%, the track was terminated.

2. Condition 2. Effects of notched noise level and signal frequency

The bandwidth of the test stimuli was set at 800 Hz and f_h was either 1.2 or 4 kHz, based on the simulations reported in the previous section. The modulation frequency was varied from 4–256 Hz in equal logarithmic steps. The spectrum level of the notched noise was set at -20 or 40 dB SPL.

3. Condition 3. Effects of signal bandwidth on the TMTF

The upper cutoff frequency of the signal was fixed at 4 kHz as mentioned previously, so that as the bandwidth increased the signal would not extend to higher frequency regions where peripheral filters would be broader. Bandwidths were 200, 400, 800, and 1600 Hz. Thresholds were also measured with a broadband carrier with an upper cutoff frequency of 10 kHz. The modulation frequencies were the same as those used in condition 2, with the limitation that modulation frequency could not exceed half the bandwidth of the test stimulus. The notched noise spectrum level was set at 40 dB SPL. No notched noise was used with the broadband carrier.

E. Results

1. Effects of notched noise level and signal frequency on the TMTF

a. Effects of notched noise level. Figure 5 shows data for $f_h = 1.2$ kHz (left column) and 4 kHz (right column), with the four subjects shown across rows, and notched noise level
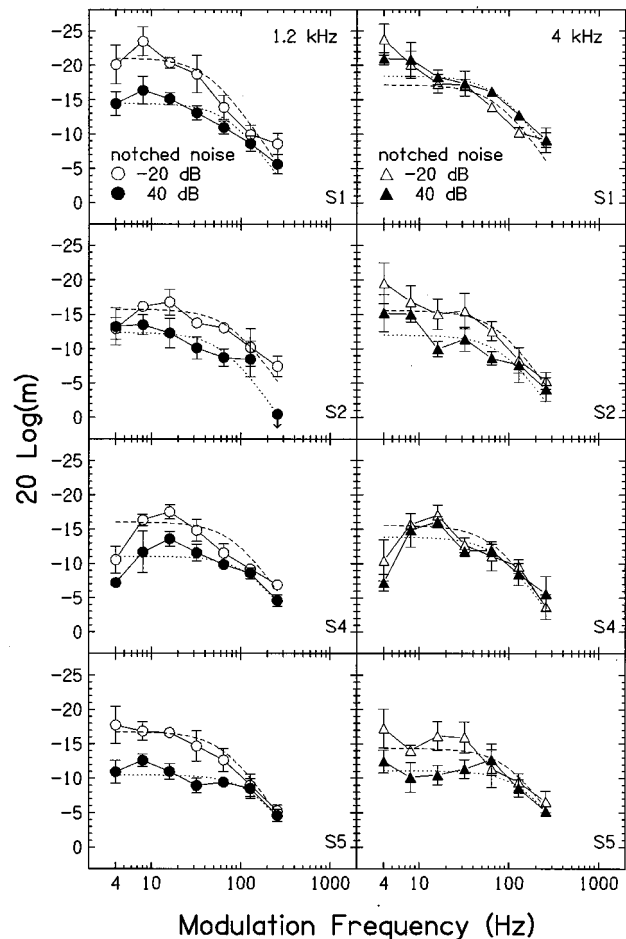
FIG. 5. Thresholds for f_h of 1.2 and 4 kHz for four subjects. The spectrum level of the notched noise is -20 or 40 dB SPL. The dashed lines are first-order Butterworth filter functions fit by a least-squares criterion.

as the parameter. The bandwidth is 800 Hz. Thresholds are shown in the presence of notched noise at a spectrum level of -20 (open symbols) and 40 dB (filled symbols). The symbol with the downward arrow (column 1, S2) indicates that no threshold could be obtained in that condition, due to termination of the track. For all subjects, increasing the level of the notched noise affected the form of the TMTF for the 1.2-kHz region (left column). Increasing the spectrum level of the notched noise from -20 to 40 dB increased thresholds by approximately 5 dB for the lowest modulation frequencies, but had less effect as f_m increased. For two subjects (S2 and S5) notched noise had the same effect on the TMTF for the 4-kHz region, while the other two subjects showed no significant effect.

Data points from f_m of 8–256 Hz were fit with a first-order lowpass Butterworth filter by a least-squares algorithm, as shown by the dashed lines. The S, F_c, and $\tau(1/2\pi F_c)$ calculated from these fits are presented in Table I. The average rms error was 1.29 dB. Although this function rolls off at 6 dB/oct rather than the 3 dB/oct that has been seen in broadband TMTF data (Viemeister, 1979; Forrest and Green, 1987), it is a realizable lowpass filter function, and the time constant of the impulse response is meaningfully related to the cutoff frequency.

TABLE I. Sensitivity (S), cutoff frequency (F_c) and the time constant (τ) obtained from lowpass filter functions fit to the data in Fig. 5.

f_h (kHz)	Noise level (dB SPL)	Subject	S	F_c (Hz)	τ (ms)
1.2	−20	S1	21.10	44.21	3.6
		S2	15.77	80.28	2.0
		S4	16.06	84.91	1.9
		S5	16.77	66.91	2.4
		mean	16.46	68.26	2.3
	40	S1	14.48	79.54	2.0
		S2	12.39	66.68	2.4
		S4	11.08	131.60	1.2
		S5	10.50	143.90	1.1
		mean	11.33	136.30	1.2
4	−20	S1	17.16	75.4	2.1
		S2	15.57	76.08	2.1
		S4	15.59	68.06	2.3
		S5	14.39	101.30	1.6
		mean	16.35	64.89	2.5
	40	S1	18.45	78.82	2.0
		S2	12.02	88.23	1.8
		S4	13.87	79.72	2.0
		S5	11.14	150.10	1.1
		mean	13.73	109.30	1.5

III. PREDICTIONS USING THE ENVELOPE DETECTOR

A. Signal frequency region varied

One of the aims of this study is to quantify the expected effects of peripheral filtering on temporal resolution. Therefore, simulations were performed in order to compare the data to predictions of the envelope detector when the initial stage was a fourth-order gammatone filter. As in the preliminary simulations, the ERB of the gammatone filter was set to equal the ERB of an auditory filter measured by the notched-noise method using the formula proposed by Glasberg and Moore (1990). The signal was then half-wave rectified and filtered by a first-order lowpass Butterworth. The decision statistic was the fourth moment of the envelope, a measure of noise power fluctuation (Hartmann and Pumplin, 1988). This statistic was chosen because it has been used to predict modulation detection and modulation masking data (Strickland and Viemeister, 1996), and could predict thresholds as low as those seen in the data. The cutoff frequency of the lowpass filter was set at 250 Hz, based on the best fit of the model with the fourth moment decision statistic to broadband TMTF data in a previous study (Strickland and Viemeister, 1996).

The stimulus parameters were the same as those in condition 2 for the higher notched noise level, that is a band-

Although most of the TMTFs are lowpass functions, for some subjects (particularly S4) and conditions there is a distinct highpass segment at low modulation frequencies. This highpass segment is typical of TMTFs measured with gated carriers, particularly at short durations (see, e.g., Viemeister, 1979; Sheft and Yost, 1990).

b. Effects of signal frequency. The data are replotted in Fig. 6, to show the effects of frequency region. Notched noise level is shown across columns. Circles show data for f_h of 1.2 kHz, and triangles for 4 kHz. For S2, S4, and S5, TMTFs are very similar for the two different signal frequency regions. For S1 in the higher level of notched noise, the TMTF for the 1.2-kHz region has a lower S than the one for the 4-kHz region, but F_c is nearly the same.

2. Effects of signal bandwidth on the TMTF

In Fig. 7, TMTFs are shown for three subjects and their average with signal bandwidth as the parameter. The upper cutoff frequency of the signal was fixed at 4 kHz, and bandwidth was 200, 400, 800, or 1600 Hz. Only modulation frequencies less than half the bandwidth were measured for each bandwidth condition. The S for each function was defined as the absolute value of the average of thresholds for modulation frequencies of 4, 8, and 16 Hz. The S increased by 3–4 dB for each doubling of bandwidth. In Fig. 8, the functions are normalized to the S for the 1600-Hz bandwidth for each subject. The functions nearly superimpose, indicating that F_c does not measurably decrease with bandwidth.

FIG. 6. The data from Fig. 5 replotted so that thresholds for test stimuli with f_h of 1.2 and 4 kHz may be compared within panels.
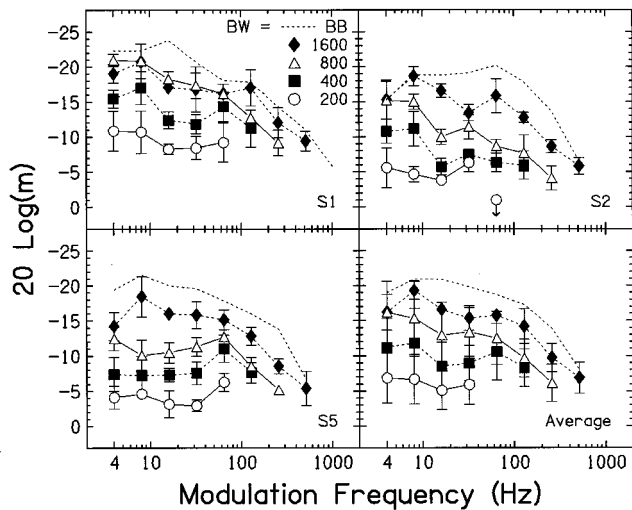
FIG. 7. Thresholds for a test stimulus with f_h of 4 kHz, and a bandwidth of 200, 400, 800, or 1600 Hz, for three subjects and their average. The dashed lines with no symbols show modulation thresholds for a broadband noise carrier.

width of 800 Hz, f_h of 1.2 and 4 kHz, and notched noise at a spectrum level 10 dB below that of the signal. The optimal gammatone filter placements for the two signal frequencies were at center frequencies of 1 and 3.6 kHz. For each trial, the difference in the fourth moment for the signal and the nonsignal intervals was measured. The mean and standard deviation of 50 trials was used to estimate d'. Thresholds were estimated by fitting a least squares line to $\log(d')$ vs $\log(m)$ and estimating the modulation depth for a d' of 1.285, which is the value tracked in the 3IFC task used to collect the data.

Figure 9 shows the thresholds predicted by the model (upper left panel), as well as the average thresholds for the four subjects (bottom right panel). The model predicts that both the S and the F_c of the TMTF will be lower for the 1.2-kHz region than for the 4-kHz region. Both effects are due to the difference in bandwidth between the two gamma-
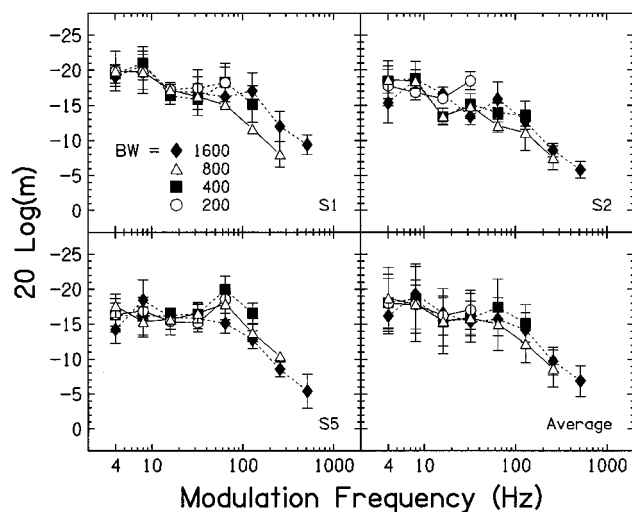
FIG. 8. The data from Fig. 7 normalized to the low-frequency plateau for the 1600-Hz bandwidth for each subject.
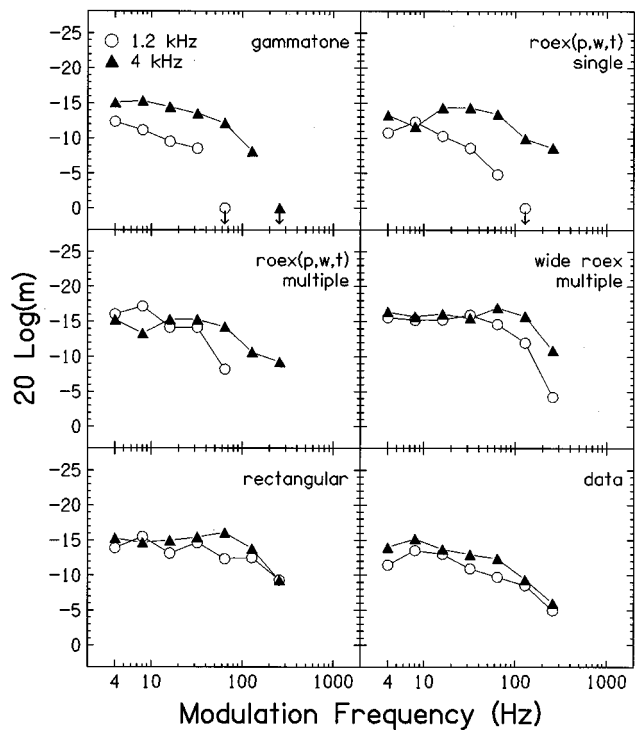
FIG. 9. Data (bottom right panel) and predictions of the envelope detector (all other panels) for test stimuli with f_h of 1.2 or 4 kHz in notched noise with a spectrum level 10 dB below that of the test stimulus. The decision statistic is the fourth moment of the envelope. See text for an explanation of the model parameters.

tone filters. The equivalent rectangular bandwidth (ERB) of the gammatone filter centered at 1 kHz is 133 Hz, while the ERB of the gammatone filter centered at 3.6 kHz is 413 Hz. In contrast, the data from the subjects show little change in S or F_c with frequency region. This suggests that peripheral filtering does not limit temporal resolution under the present conditions.

A possible explanation for the discrepancy between the predictions of the model and the data is that the gammatone filters used as the first stage may underestimate auditory filter widths, particularly because of the fairly high levels of the test stimuli. In a recent study, Rosen et al. (1992) argue that the change in auditory filter shape with level is not determined by the masker (input) level, as had been previously assumed, but rather by the probe (output) level. Filter bandwidths measured with output level constant increase much more with level than those measured with input level constant. The gammatone filters used in the simulations were fit to filter shapes estimated on the basis of input level, and therefore may have underestimated auditory filter widths. Therefore, new simulations were run in which thresholds were predicted by the model when the first stage was one or more roex (p,w,t) filters measured by Rosen et al. (1995). Their filter shapes for center frequencies of 1 and 4 kHz were extrapolated to 1.08 and 3.9 kHz for the lower and higher frequency regions, respectively. The level parameters for the two filters were set at 70 and 76 dB, based on the approximate level of the test stimuli that would be passed by the filters. Results of simulations using a single roex (p,w,t)

filter are shown in the upper right panel of Fig. 9. The S and F_c are still lower for the lower frequency region. The filter centered at 3.8 kHz passed most of the test stimulus, while the narrower filter at 1 kHz passed less than half of the test stimulus bandwidth. To make the simulation more realistic, thresholds were predicted when information from all of the filters in the test stimulus bandwidth were combined by taking the square root of the sum of the squared d'''s. Filter center frequencies were chosen so that adjacent filters overlapped at the half-power point, in an attempt to keep the output of the filters independent. Filter CFs were 0.46, 0.57, 0.71, 0.88, and 1.08 kHz for the lower frequency signal, and 3.4 and 3.85 kHz for the higher frequency signal. Results of these simulations are shown in the middle left panel of Fig. 9. The S is nearly equal for the two frequency regions, and the thresholds for the higher frequency region are quite close to those found in the data, but the model still predicts a lower F_c for the lower frequency region.

Simulations were done to determine how wide the peripheral filters would have to be to predict the results. The roex (p,w,t) filter shapes were used, but the parameters were adjusted to double the 3-dB bandwidths. It was estimated that this would produce approximately the desired F_c for the lower frequency region. For the 4-kHz frequency region, two filters centered at 3.2 and 4 kHz were used, based on the half-power criteria used above. For the 1.2-kHz region, filters were centered at 0.43, 0.58, 0.78, and 1.05 kHz. Results of these simulations are presented in the middle right panel of Fig. 9. The thresholds for the lower frequency region are now predicted fairly well, but the predicted TMTF for the higher frequency region has an F_c that is higher than that in the data. Thus the model predicts the data fairly well with roex filters in the higher frequency region, and broadened roex filters in the lower frequency region.

A limiting, although probably unrealistic, condition is presented in the lower left panel. In this case the first stage of the model is a single matched rectangular filter. This would be consistent with listeners combining amplitude and phase information across multiple peripheral channels. The predictions match the data fairly well.

B. Signal bandwidth varied

Because roex filters allowed the model to successfully predict the 4-kHz data, they were used to predict the effects of signal bandwidth. As in condition 3, f_h was fixed at 4 kHz, and test stimulus bandwidths were 200, 400, 800, or 1600 Hz. The notched noise level was 10 dB below that of the test stimulus. For the 200- and 400-Hz bandwidths, a roex filter centered at 3.9 kHz was used. For the 800-Hz bandwidth, filters were centered at 3.4 and 3.85 kHz. Filters centered at 2.6, 3, 3.4, and 3.9 kHz were used for the 1600-Hz bandwidth. Thresholds are presented in Fig. 10, along with the average data for the three subjects. Notice that F_c, insofar as it may be measured, does not change with test stimulus bandwidth for the predicted data nor the average data. Also, for both the model and the data, S increases with bandwidth.
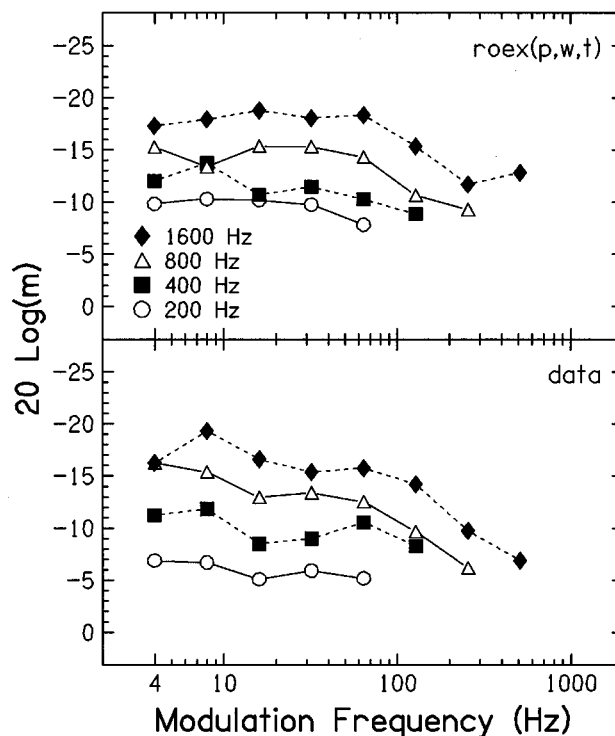
FIG. 10. The top panel shows thresholds predicted by the envelope detector when multiple roex (p,w,t) filters are the first stage, and the decision statistic is the fourth moment. The bottom panel shows average subject data replotted from Fig. 7.

IV. DISCUSSION

A. Effects of notched noise

As set forth in the Introduction, notched noise might be expected to affect the TMTF for two reasons. One is that listeners might be able to attend to broader high-frequency channels in quiet, and the presence of notched noise would force them to use lower-frequency channels which would have narrower bandwidths. The other is that listeners might be able to use nonlinear upward spread of excitation in quiet, and this would be prevented by the notched noise masker. The results are more consistent with the latter hypothesis than the former. Note that notched noise had an effect only for the lower modulation frequencies. If listeners were originally attending to broader, high-frequency filters, and were then forced to listen to narrower, low-frequency filters, the effects should be greatest at the highest modulation frequencies. But in fact the data show almost no effect of the notched noise at the highest modulation frequencies.

If listeners attend to the frequency region above the signal, the effective modulation depth may be increased by the nonlinear upward spread of excitation. As noted in the Introduction, previous research (Zwicker, 1976; Nelson, 1992) suggests that this occurs for modulation frequencies at or below 64 Hz. This is consistent with the results of this study (see Fig. 5). Increasing the level of the notched noise increased thresholds for modulation frequencies at or below 64 Hz but did not affect thresholds for higher modulation frequencies, suggesting that it was eliminating off-frequency listening.

The consequence of this dynamic component of modulation detection is that off-frequency listening does not change the form of the TMTF in a simple manner. As suggested by the results shown in Fig. 5, off-frequency listening decreases thresholds at low modulation frequencies but not at high, so that the TMTF appears to have a higher S, but a lower F_c (see Table I); that is, off-frequency listening results in an underestimate of temporal resolution.

B. Effects of signal bandwidth

Although it was not possible to measure F_c for the narrower bandwidths, if the functions were equalized for S (Fig. 8) they largely superimposed, indicating there was no measurable change in F_c with signal bandwidth. These results differ from those of Eddins (1993), who found a clear increase in cutoff frequency with bandwidth. As discussed in the Introduction, this reflects the fact that filtering after modulation, as in Eddins' study, may impose a bandwidth effect by decreasing the modulation depth of the signal as a function of bandwidth. The AM versus QFM or random noises used in this study are not filtered after modulation; therefore the modulation depth is independent of bandwidth. As shown by the simulations, the bandwidth data in the present study are consistent with the predictions of the envelope detector with multiple roex filters in the initial stage.

The cutoff frequencies measured in the present study are consistently higher than those measured by Eddins (1993). For example, in Table I, the average F_c for a signal with an f_h of 4 kHz in notched noise at a spectrum level of -20 dB SPL is 64.89 Hz, while for a comparable condition in Eddins study the average F_c is only 44.7 Hz. This is probably not solely due to the fact that he filtered his signals after modulation. The signals and notched noise in the present study were gated, while in the Eddins study the signals were gated but presented in continuous broadband noise. Average TMTFs for broadband noise are shown in Fig. 11 for a continuous carrier, from Strickland and Viemeister (1996), and a gated carrier, from the present experiment. Gating the carrier increases thresholds for low modulation frequencies but not for high. Although only one subject was common to both conditions, the gated-continuous difference is comparable to that measured by Viemeister (1979) in which the same subjects were used in both conditions. The cutoff frequencies for filters fit to the functions in Fig. 11 are 90 and 130 Hz for the continuous and gated carriers respectively. Thus gating the carrier causes an increase in F_c, but this should not be interpreted as in increase in temporal resolution: modulation thresholds at high modulation frequencies, where temporal resolution affects performance, are unaffected by gating.

In this study S increased by 3–4 dB for each doubling of bandwidth. This increase is greater than that reported in several other studies. Figure 12 shows S as a function of signal bandwidth from this study compared to S from studies by Maiwald (1967), Eddins (1993), and Viemeister (1979). In this figure, S is the absolute value of the average of thresholds for modulation frequencies of 4, 8, and 16 Hz, except for the data of Maiwald, where S is the absolute value of the threshold at a modulation frequency of 4 Hz. Although the
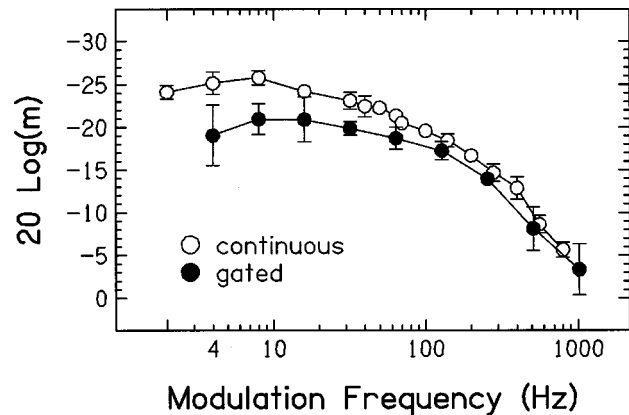
FIG. 11. Average broadband TMTFs for a continuous carrier and a gated carrier.

slope of the function relating threshold to bandwidth is comparable for the other three studies, it appears to be steeper for the present data. At these low modulation frequencies, effects of filtering before or after modulation should not affect thresholds. The most likely reason for the difference is the fact that the signals in this study were surrounded by notched noise at a signal-to-noise ratio of 10 dB, while in the other studies there was either no surrounding noise (Maiwald, 1967; Viemeister, 1979) or a lower level of broadband noise (Eddins, 1993).

An analogous result has been reported for intensity discrimination of narrowband noises. When the signal bandwidth is narrower than the pedestal bandwidth, the slope of the function relating threshold in $10 \log(\Delta I/I)$ to log (bandwidth) is steeper than when signal and pedestal bandwidths are equal (Schacknow and Raab, 1976). Specifically, at narrow bandwidths, thresholds are lower when the pedestal bandwidth is equal to the signal bandwidth than when the pedestal is greater than signal bandwidth. The present stimulus configuration is analogous to the condition where the pedestal bandwidth is greater than the signal bandwidth, in that there is essentially a broadband noise, of which only a
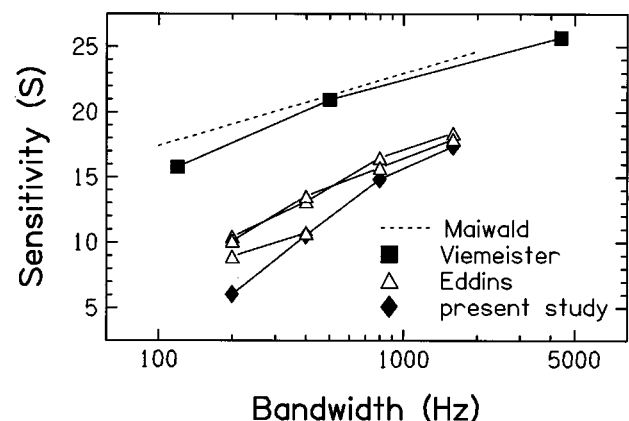
FIG. 12. S as a function of bandwidth for four studies: Maiwald (1967), Viemeister (1979), Eddins (1993), and the present study. Data points are the absolute value of the average threshold for modulation frequencies of 4, 8, and 16 Hz, except for those of Maiwald, which are for 4 Hz only.

narrow band is modulated. The conditions used by Maiwald, Viemeister, and Eddins were analogous to intensity discrimination with signal bandwidth equal to pedestal bandwidth. There are two factors that might explain why thresholds increase more rapidly as bandwidth decreases. The greater increase in thresholds with decrease in bandwidth seen in this study and in the Schacknow and Raab study could be due to notched noise leaking into a peripheral filter. This would be supported by the fact that the rate of increase in threshold is less between the bandwidth change from 1600 to 800 Hz, which would presumably be greater than the bandwidth of a filter centered on the signal, than the rate of increase as bandwidth is decreased below 800 Hz. The other possibility is that upward spread of excitation is more usable for narrower bandwidths, therefore adding notched noise around the signal increases thresholds more for narrow bandwidths than for broader ones.

C. Explanation for the lack of change in the TMTF with signal frequency

When the initial simulations were done to choose the signal conditions, gammatone filters with ERBs measured by the notched-noise method using the formula proposed by Glasberg and Moore (1990) were used as estimates of peripheral filtering. Thresholds predicted with these filters roll off more sharply than the data. Filter shapes estimated by Rosen *et al.* (1992, 1995), which widen and become increasingly low pass with level, lead to good predictions for the higher frequency region, but not for the lower frequency region. In the lower frequency region, the data are well predicted only if filter bandwidths are approximately doubled. In the envelope detector model, if the initial bandpass filters have large bandwidths, the cutoff frequency of the lowpass filter will determine the cutoff frequency of the TMTF. Above moderate signal levels, the width of the initial filters may be wide enough that the final filter may be the determining factor for all but very low signal frequencies.

V. SUMMARY AND CONCLUSIONS

Resolution of envelopes by the auditory system was studied by examining the effects of changing the narrowband noise frequency and bandwidth, and of restricting the listening region. The results were interpreted using the envelope detector model.

Restricting the listening region with notched noise increased thresholds for low modulation frequencies but not for high. This is not understandable in terms of the simple linear envelope detector, but is consistent with other data showing that upward spread of excitation may be ''sluggish,'' and may magnify modulation depth only for low modulation frequencies.

The TMTFs measured in frequency regions of 1.2 and 4 kHz were nearly identical. Because auditory filter widths change substantially between these frequencies, these results indicate that for the frequencies and levels used in our experiments, auditory temporal resolution is not limited by auditory filtering. It appears that the auditory system can adjust its effective bandwidth depending on the task. When spectral resolution is important, as in most measurements of auditory filtering, a narrow filter is used. When temporal resolution is important, as in the present study, the effective filter bandwidth is broad. One possibility is that a broad effective bandwidth is achieved by combining the outputs of narrow filters with different center frequencies. Such combination would almost certainly occur central to the auditory nerve and thus, at least indirectly, would involve pooling of spike trains. As discussed by Viemeister (1979), accurate recovery of high-frequency envelopes requires preservation of amplitudes and phases with a degree of fidelity that may be incompatible with data from the auditory nerve. Another possibility is that peripheral filtering does not limit temporal resolution, and that the filtering measured psychophysically reflects ''sharpening'' by central mechanisms. At present, neither the physiological nor the psychophysical data permit strong conclusions about which of these possibilities is more correct.

[1]Goldstein (1967) has shown that for a SAM tone the greatest modulation depth, measured as maximum-to-minimum ratio, will be obtained by a bandpass filter that is placed so that one sideband is greatly attenuated, and the level of the carrier is attenuated to the level of the remaining sideband. This ''beat'' envelope will not be sinusoidal, but will beat at a rate equal to the difference in frequency between the carrier and the sideband. For the SAM noises used in this study, in which the sidebands extend out from the carrier band by the amount of the modulation frequency (see Fig. 3), this filter placement may be achieved at high modulation frequencies. At these high modulation frequencies, as the modulation index of the signal is decreased, and thus the sideband level is decreased, the modulation depth at the output of the filter may be kept constant by shifting its center frequency away from the carrier band so that it is still attenuated to the level of the sideband. This yields a TMTF that does not roll off as expected at high modulation frequencies, because performance should not decrease with decreasing modulation index. This condition does not arise when QFM is the comparison, because the filter placement that will produce beats in the AM noise will also produce beats in the QFM noise, and thus will not be optimal for discriminating between the two. The filter must be placed so that it will pass both sidebands, resulting in a TMTF that rolls off in the expected manner.

Arndt, S. C. (**1993**). ''Masking at envelope peaks and envelope valleys in amplitude modulated tonal maskers,'' Master's thesis, University of Minnesota.

Buunen, T. J. F. (**1976**). ''On the perception of phase differences in acoustic signals,'' Ph.D. dissertation, Technische Hogeschool, Delft.

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.

Fassel, R., and Kohlrausch, A. (**1996**). ''Sinusoidal amplitude modulation thresholds as a function of carrier frequency and level,'' J. Acoust. Soc. Am. **99**, 2566(A).

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.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hearing Res. **47**, 103–138.

Goldstein, J. L. (**1967**). ''Auditory spectral filtering and monaural phase perception,'' J. Acoust. Soc. Am. **41**, 458–479.

Hartmann, W. M., and Pumplin, J. (**1988**). ''Noise power fluctuations and the masking of sine signals,'' J. Acoust. Soc. Am. **83**, 2277–2289.

Jeffress, L. A. (**1967**). ''Stimulus-oriented approach to detection re-examined,'' J. Acoust. Soc. Am. **41**, 480–488.

Joris, P. X., and Yin, T. C. T. (**1992**). ''Responses to amplitude-modulated tones in the auditory nerve of the cat,'' J. Acoust. Soc. Am. **91**, 215–232.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Maiwald, D. (**1967**). ''The calculation of modulation thresholds with a model,'' Acustica **18**, 193–207.

Nelson, D. A. (**1992**). ''Temporal resolution within the upper ''accessory excitation'' region for maskers with fluctuating envelopes,'' J. Acoust. Soc. Am. **91**, 2356(A).

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), pp. 123–177.

Patterson, R., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (**1987**). ''An efficient auditory filterbank based on the gammatone function,'' Paper presented at a meeting of the IOC Speech Group on Auditory Modelling at RSRE, 14–15 December 1987.

Rice, S. O. (**1944**). ''Mathematical analysis of random noise,'' Bell Syst. Tech. J. **23**, 282–332; **24**, 46–156 (1945). Also available in N. Wax (ed.), *Selected Papers on Noise and Stochastic Processes* (Dover, New York, 1954).

Rodenburg, M. (**1977**). ''Sensitivity of the auditory system to differences in intensity,'' Ph.D. dissertation, Erasmus University of Rotterdam, Netherlands.

Rosen, S., Baker, R. J., and Darling, A. M. (**1995**). ''Nonlinearities of human auditory filtering,'' in Abstracts of the Eighteenth Annual Midwinter Research Meeting of the Association for Research in Otolaryngology, 5–9 February 1995, St. Petersburg Beach, FL, Abstract Number 214, p. 54.

Rosen, S., Baker, R. J., and Kramer, S. (**1992**). ''Characterizing changes in auditory filter bandwidth as a function of level,'' Adv. Biosci. **83**, 171–177.

Schacknow, P. N., and Raab, D. H. (**1976**). ''Noise-intensity discrimination: Effects of bandwidth conditions and mode of masker presentation,'' J. Acoust. Soc. Am. **60**, 893–905.

Sheft, S., and Yost, W. A. (**1990**). ''Temporal integration in amplitude modulation detection,'' J. Acoust. Soc. Am. **88**, 796–805.

Strickland, E. A., and Viemeister, N. F. (**1996**). ''Cues for discrimination of envelopes,'' J. Acoust. Soc. Am. **99**, 3638–3646.

van Zanten, G. A. (**1980**). ''Temporal modulation transfer functions for intensity modulated noise bands,'' in *Psychophysical, Physiological and Behavioural Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft U.P., Delft, The Netherlands), pp. 206–209.

Viemeister, N. F. (**1977**). ''Temporal factors in audition: A systems analysis approach,'' in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 419–427.

Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based upon modulation thresholds,'' J. Acoust. Soc. Am. **66**, 1364–1380.

Zwicker, E. (**1976**). ''Masking period patterns of amplitude modulated pure tones,'' Acustica **36**, 113–120.

A unitary model of pitch perception

Ray Meddis and Lowel O'Mard

Department of Psychology, Essex University, Colchester CO4 3SQ, United Kingdom

A model of the mechanism of residue pitch perception is revisited. It is evaluated in the context of some new empirical results, and it is proposed that the model is able to reconcile a number of differing approaches in the history of theories of pitch perception. The model consists of four sequential processing stages: peripheral frequency selectivity, within-channel half-wave rectification and low-pass filtering, within-channel periodicity extraction, and cross-channel aggregation of the output. The pitch percept is represented by the aggregated periodicity function. Using autocorrelation as the periodicity extraction method and the summary autocorrelation function (SACF) as the method for representing pitch information, it is shown that the model can simulate new experimental results that show how the quality of the pitch percept is influenced by the resolvability of the harmonic components of the stimulus complex. These include: (i) the pitch of harmonic stimuli whose components alternate in phase; (ii) the increased frequency difference limen of tones consisting of higher harmonics; and (iii) the influence of a mistuned harmonic on the pitch of the complex as a function of its harmonic number. To accommodate these paradigms, it was necessary to compare stimuli along the length of the SACF rather than relying upon the highest peak alone. These new results demonstrate that the model responds differently to complexes consisting of low and high harmonics. As a consequence, it is not necessary to postulate two separate mechanisms to explain different pitch percepts associated with resolved and unresolved harmonics. © *1997 Acoustical Society of America.* [S0001-4966(97)06509-0]

PACS numbers: 43.66.Nm, 43.66.Ba, 43.66.Hg [RHD]

INTRODUCTION

The study of virtual (or low) pitch perception has a long history of controversy. Theories have typically championed either spectral (e.g. Terhardt, 1979; Goldstein, 1973) or temporal (e.g., Schouten, 1970; Licklider, 1951, 1959) explanations. Recently Houtsma and Smurzynski (1990) have shown that the pitch of stimuli consisting of high-numbered harmonics may be qualitatively different from and weaker than stimuli consisting of low-numbered harmonics. Patterson (1987) has also shown that listeners are better at detecting component phase shifts in such complexes when the harmonic numbers are higher. These and other results suggest that there may be two mechanisms for extracting pitch, one for low harmonics and one for high harmonics. Houtsma and Smurzynski (1990) give a useful discussion of this issue. Since spectral theories typically emphasize results based on low-numbered harmonics and temporal theories typically emphasize studies using high-numbered harmonics, this might be interpreted to mean that a spectral mechanism generates pitch percepts for low-frequency stimuli while a temporal mechanism generates pitch percepts for high-frequency stimuli. Carlyon and Shackleton (1994) have explicitly endorsed this view. The following investigations show, however, that both types of result can be embraced by a single model. This development offers the possibility that the traditional opposition of spectral and temporal approaches might be resolved in a single system.

We recently presented a computational model of pitch perception based on autocorrelation of the estimated probabilities of firing in groups of auditory nerve fibers (Meddis and Hewitt, 1991a, b). Although autocorrelation had been used before in this context (e.g., Licklider, 1959; Lyon, 1984), the novelty of our system lay in the use of a summary autocorrelation function (SACF) as the basis of the prediction of the pitch percepts of human listeners (see also Assmann and Summerfield, 1990). The SACF is the vector sum across all audible frequencies of the autocorrelation functions (ACF) of auditory nerve-fiber firing probabilities. The idea of summarizing temporal information across channels is a characteristic of a number of other ''temporal'' models (van Noorden, 1982; Moore, 1977; Patterson, 1987).

Meddis and Hewitt's (1991a) model can be thought of as a special case of a class of a more general model consisting of four stages (see Fig. 1);

(1) peripheral (mechanical) band-pass filtering,
(2) half-wave rectification and low-pass filtering,
(3) within-channel periodicity extraction,
(4) across-channel aggregation of periodicity estimates.

Pitch decision algorithms are then applied exclusively to the results of the cross-channel aggregation in stage 4.

Meddis and Hewitt (1991a, b) used autocorrelation as one possible way of approximating the within-channel periodicity extraction stage. It was shown that the model was able to predict listeners' performance in the following areas; pitch of the missing fundamental, ambiguous pitch, pitch shift of equally spaced inharmonic components, musical chords, repetition pitch, the pitch of interrupted noise, the existence region, and the dominance region for pitch. The model was also able to simulate a number of aspects of listeners' sensitivity to the phase relationships among adjacent harmonic components of tone complexes. This phase sensi-

(ACF)

periodicity detection

bandpass filters

half-wave rectification low-pass

ΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣ
cross-channel summation

(SACF)
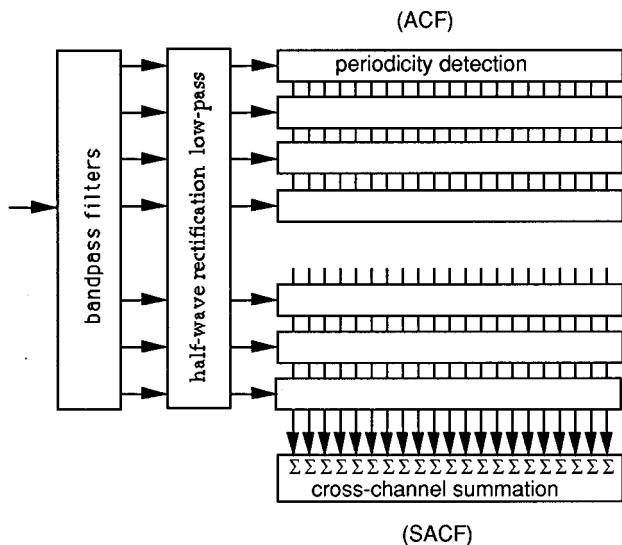
FIG. 1. General scheme of the four-stage pitch extraction model.

tivity was successfully demonstrated in the model using (i) amplitude-modulated and quasi-frequency-modulated stimuli, (ii) harmonic complexes with alternating phase and monotonic phase change across harmonic components, and (iii) phase effects associated with mistuned harmonics.

Such models have value in at least three different domains. First, they contribute to the development of psychological theories of the nature of pitch perception. Second, they serve as a class of computing engines acting as components of general models of auditory scene analysis (e.g., Brown and Cooke, 1994). Third, they offer a potential guide to physiological studies of the neurological mechanisms underlying pitch perception. Stages 1 and 2 of the model are consciously modeled on middle- and inner-ear functioning. Stages 3 and 4 are presumed to occur in the auditory brainstem although detailed theories of this process have yet to emerge. Our use of autocorrelation to simulate stage 3 of the model may appear unlikely as a physiological mechanism and the topic will be reviewed in the discussion.

In this report, we shall revisit the model by looking at three crucial phenomena not modeled before; (i) the pitch of alternating-phase harmonic complexes, (ii) frequency-difference limens for virtual pitch, and (iii) the effect of mistuning individual harmonics on the pitch of the whole complex. The phenomena are crucial in the sense that they could be used to justify dual (temporal/spectral) mechanisms of pitch perception. The presentation will emphasize an emergent property of the model whereby it responds to resolved and unresolved harmonics in a qualitatively different way without the need to postulate additional mechanisms.

I. MODEL DESCRIPTION

Full details of the model are given in Meddis and Hewitt (1991a). It has not been altered in any substantial way. Some details of the implementation of the individual stages are given below and a list of all parameters is given in the appendix. The model was implemented using version 1.9.2 of the LUTEar computational library.[1] An example of the output of the model is given in Fig. 2 showing the ACFs and the

SACF in response to a 100-Hz fundamental tone consisting of groups of low- and high-numbered harmonics. The model was evaluated using a sampling interval of 50 μs. All parameters are summarized in the Appendix.

Stimulus generation. The stimuli used were approximately 100 ms long with an onset (raised sine) ramp of 2.5 ms. The stimulus duration was always adjusted to be an integer multiple of the period of the signal. This is because the SACF pulsates in synchrony with the stimulus and the snapshot is best taken at a fixed point in the cycle. All tone components were harmonic and presented at 60 dB SPL unless otherwise stated.

Pre-emphasis filter. This filter is a simplified representation of the pressure gain of the middle ear. It is implemented using a second-order band-pass filter with skirts down by 3 dB at 450 and 8500 Hz.

Band-pass filters (stage 1). A bank of 60 linear fourth-order gammatone filters (Patterson *et al.*, 1992) was used to represent the mechanical frequency selectivity of the cochlea. Filter center frequencies were equally spaced on a log scale between 100 and 8000 Hz. The bandwidths of the filters were chosen using the formula suggested by Glasberg and Moore (1990).

Hair cell (stage 2). The output from each filter was passed to an inner hair cell model (Meddis, 1986, 1988), representing mechanical to neural transduction. [The parameters are the same as those given in Meddis *et al.* (1990) except for parameters A and B which were changed to 100 and 6000, respectively, to accommodate a change from arbitrary units to units of pressure, μPascals.] The output from the simulated hair cell was in the form of a continuously changing probability of an action potential, $p(t)$, in a postsynaptic auditory nerve fiber. The output from the hair cell stage is characterised by an ac and a dc component. The ac component has the same frequency as the signal but is increasingly attenuated above 1 kHz. The dc component is a saturating monotonic function of signal level. Refractory effects were not modeled.

Periodicity detection (stage 3). A running autocorrelation function was computed in each channel based on the firing probabilities. The reported SACF is a ''snapshot'' of the function at the time of the cessation of the stimulus. More strictly, the function is the running cross product of the probabilities at a lag l;

$$h(t,l,k) = \frac{1}{\tau} \sum_{i=1}^{\infty} p(t-T,k)(t-T-l,k)e^{(-T/\tau)} \, dt, \quad (1)$$

where $h(t,l,k)$ is the ACF at lag l, time t, in channel k, $T = idt$, τ is the time constant (normally 10 ms), and dt is the sampling interval.

Cross-channel summation (stage 4). All ACFs were summed with equal weight to create the SACF $s(t,l)$,

$$s(t,l) = \sum_{k=1}^{N} h(t,l,k), \quad (2)$$

where N is the number of channels used.

R. Meddis and L. O'Mard: Unitary model of pitch perception
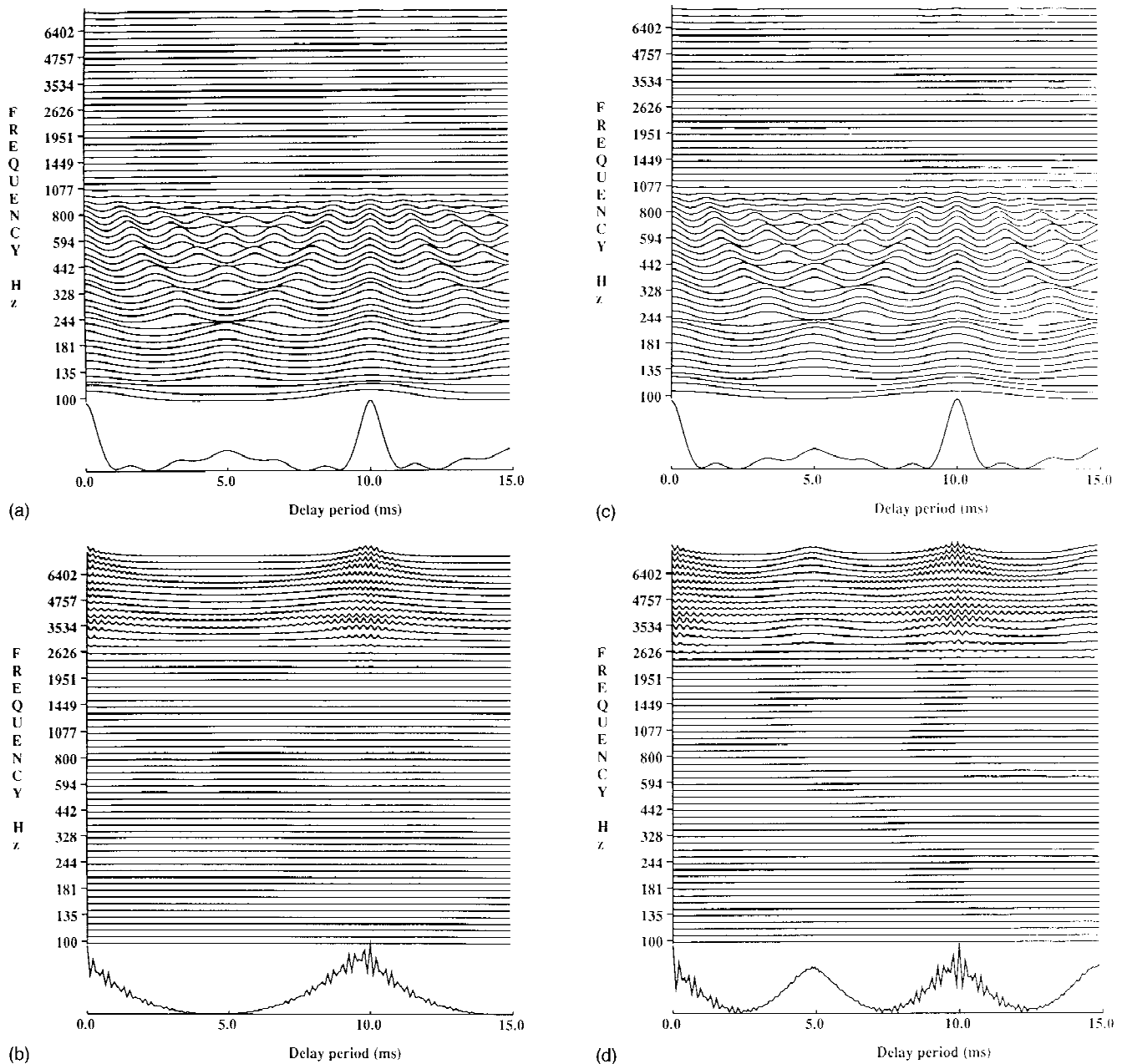
FIG. 2. Autocorrelation functions (ACF) and summary autocorrelation function (SACF) for harmonic complexes with 100-Hz fundamental. (a) and (b) Contrast results obtained using LOW (125–625 Hz) and HIGH (3900–5400 Hz) harmonics in *sine* phase. (c) and (d) Contrast results obtained using LOW and HIGH harmonics in *alternating* phase.

Pitch matching is implemented by generating the SACF for two tones and computing the squared Euclidean distance, D^2, between the two SACFs ($s(t,l)$ and $s'(t,l)$).

$$D^2 = \sum_{i=a}^{b} [s(t,l) - s'(t,l)]^2, \tag{3}$$

where $l = i \cdot dt$, and a,b defines the range of lags used in the comparison. Full details of this procedure are given in Meddis and Hewitt (1991b). A small D^2 indicates that the two stimuli are similar while a large D^2 indicates dissimilarity.

D^2 is measured over a limited segment of the SACF ($a.dt, b.dt$) and the limits of this segment are explicitly specified on each relevant occasion in the text. The segment typically covers only those lags which might reasonably in-

fluence the pitch judgments, i.e., whose reciprocals represent frequencies at which virtual pitch can be heard. As a general rule, the *qualitative* effects to be described can be observed using a wide range of different segment lengths, so long as they include lags corresponding to the pitch period. To verify this effect, the model is routinely checked by using a different SACF ranges.

Threshold judgments are simulated in a *relative* sense by establishing D^2 between a test stimulus and a comparison stimulus that differs by a fixed percentage along the dimension of interest (fundamental frequency or frequency of a mistuned harmonic). If the distance measure is greater in one situation than another then the former is judged to be more discriminable and to indicate a lower threshold (see below).

The test and comparison stimuli are chosen as typical of a threshold difference for human listeners. An alternative method used below is to define a reasonable criterion value of D^2 and then to vary the fundamental frequency until this is matched. Both methods are illustrated.

In the case of *pitch matching*, D^2 is computed between the test stimulus and a number of candidate comparison stimuli to be described below. The best match is taken to be the comparison stimulus yielding the lowest value for D^2.

The computer code and the associated parameter files for running the model are available from the authors.

II. EVALUATION OF THE MODEL

A. Pitch of alternating phase harmonics

Carlyon and Shackleton (1994) presented listeners with harmonic complexes whose component harmonics were either in sine phase or alternating phase (sine/cosine/sine/cosine.., etc.). An alternating-phase stimulus is sometimes reported as having a pitch an octave above that of its corresponding sine-phase stimulus. It can be seen in Fig. 3 that the stimulus envelope changes when the phases are alternated; the period of the envelope is only half as long. If the pitch of an alternating-phase stimulus is reported to be double that of the corresponding sine-phase stimulus, the observation could be used as evidence in favor of models of pitch extraction which depend upon the period of the envelope (temporal theories). Spectral theories, on the other hand, cannot easily explain the rise in pitch because the spectral components of the stimulus have not changed.

Carlyon and Shackleton (1994) reported that the pitch appears to double for alternating-phase stimuli when they are restricted to a small number of *high* harmonics. On the other hand, if the stimuli are restricted to a small number of *low* harmonics, the effect disappears and both sine-phase and alternating-phase stimuli are reported as having the same pitch. The results appear to imply that two different pitch mechanisms are at work; one (phase insensitive) for low harmonics and the other (phase sensitive) for high harmonics. High-harmonic stimuli generate percepts consistent with temporal theories while low-harmonic stimuli generate percepts consistent with spectral theories.

By varying the fundamental frequency and the range of harmonics used in a stimulus, Carlyon and Shackleton (1994) were able to show that the critical variable was probably not the frequency of the harmonics but their ''resolvability.'' When harmonic components are well spaced relative to the width of auditory filters (Patterson and Moore, 1986), there is relatively little interaction of adjacent harmonics within a single filter and the components are said to be ''resolved.'' This result might be used to support a more specific argument that percepts based on *resolved* harmonics may be best accounted for using spectral theories, while percepts based on *unresolved* harmonics may be best accounted for using temporal theories.

The following demonstration shows that our original model can simulate the effect without further modification. To do this we used stimuli similar to those of Carlyon and Shackleton by creating alternating-phase stimuli using only

(a)

(b)
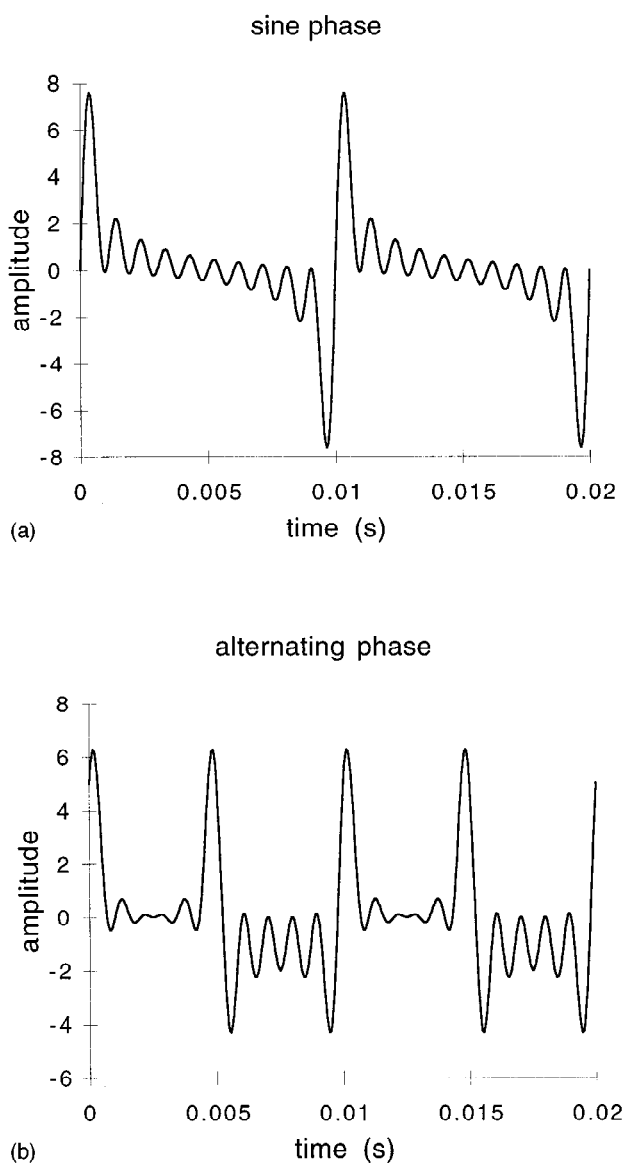
FIG. 3. Comparison of stimuli with harmonics 1–10 in (a) sine phase and (b) alternating phase. $F0=100$ Hz.

harmonics in limited-width frequency bands. The SACF for this alternating-phase stimulus was then compared with the SACFs for two similar stimuli whose harmonics were in sine phase; one with the same $F0$ and one with $F0$ set an octave higher.

The test stimulus was composed of harmonics in *alternating phase* with an $F0$ of 150 Hz. The two comparison stimuli were composed of harmonics in *sine phase* with $F0$ 150 Hz and $F0$ 300 Hz, respectively. The test was repeated three times using:

(1) only LOW harmonics (range 125–625 Hz),
(2) only MID harmonics (range 1375–1875 Hz), and
(3) only HIGH harmonics (range 3900–5400 Hz).

The stimuli were generated using a signal consisting of 80 harmonics of the fundamental, digitally filtered using a Butterworth filter set to give skirts of 24 dB per octave. The range was used to specify the 3 dB down points of the filter.

The task is to find the best match between the test stimu-

1814   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

R. Meddis and L. O'Mard: Unitary model of pitch perception   1814
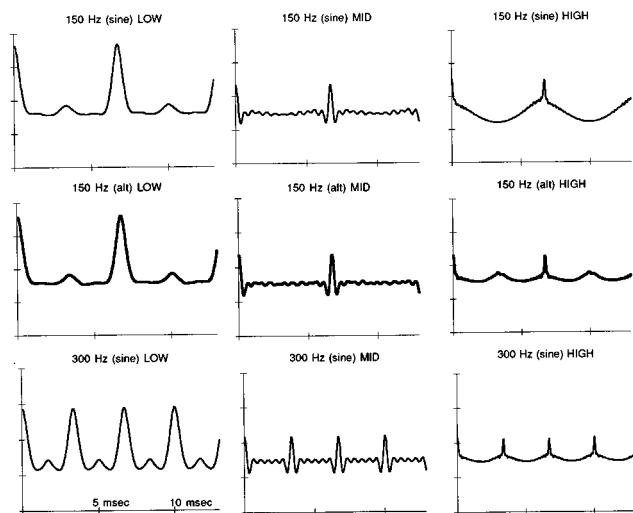
FIG. 4. The effect of alternating phase on the SACF. Top row; SACFs obtained using harmonics of 150 Hz in either LOW, MED, or HIGH ranges (see text). All harmonics are in sine phase. Middle row; as top row but the phases of successive harmonics alternate in phase. Bottom row; as top row but $F0 = 300$ Hz. Compare the SACF in the middle row (alternating phase) with the SACFs immediately above and below (sine phase). For LOW and MID, the best match is with the top row ($F0 = 150$ Hz). For HIGH, the best match is with the bottom row ($F0 = 300$ Hz).

lus (alternating phase) and either the 150-Hz or the 300-Hz comparison stimulus (sine phase). The SACFs for all nine stimuli are shown in Fig. 4, where the comparison can be made visually.

When LOW harmonics are used, the SACF of the alternating-phase 150-Hz test stimulus is indistinguishable from that of the 150-Hz sine-phase comparison stimulus. When HIGH harmonics are used, the SACF of the alternating-phase 150-Hz test stimulus generates a distinct extra peak that is missing from the sine-phase version. This makes it more similar to the SACF of the 300-Hz sine-phase comparison stimulus (bottom row). The model, therefore, predicts that a 150-Hz alternating-phase stimulus will be matched with 300-Hz sine-phase comparison stimulus when HIGH harmonics are used but not when LOW harmonics are used. This is essentially the same result as that reported by Carlyon and Shackleton (1994).

An explanation of the model's behavior can be inferred from the differences among the ACFs in Fig. 2. Figure 2(a) and (b) shows the ACFs and SACF for LOW and HIGH harmonics in *sine phase* with a fundamental frequency of 100 Hz. Contrast these figures with those given in Fig. 2(c) and (d) which is based on stimuli that are identical in all respects except that the components are in *alternating phase*. For LOW harmonics, the ACFs are indistinguishable between sine and alternating phase [2(a) and (c)]. This is because the harmonics are resolved by the filters so that the input to the individual ACFs are approximately pure tones. The ACF of a pure tone is not affected by the phase of the tone. As a consequence, changing the phase of components that are resolved by the filter bank has no effect on the pattern of ACFs and, by implication, has no effect on the percept. In the case of HIGH harmonics, on the other hand, the ACFs are different in the two corresponding figures [2(b)

and (d)]. For example, the pronounced dip at a lag of 3.3 ms for sine-phase stimuli is replaced by a distinct peak for cosine-phase stimuli. This occurs because the filters no longer resolve individual components. The within-channel waveform for multi-component signals is, therefore, strongly affected by the phase relationships of the interacting components. Autocorrelation is strongly affected by changes in shape of the waveform and this is reflected in the ACFs which predict that relative phase will be perceptually relevant when the components are unresolved (see also Patterson, 1987).

We may conclude from this demonstration that the model's behavior is consistent with the reports of listeners. Furthermore, it is not necessary to postulate two separate mechanisms to explain different percepts for resolved and unresolved harmonics when presented in alternating phase. It is a property of the four-stage model that it will respond differently to these two types of stimulus without further amendment.

The differences between the SACFs for the two stimuli are manifest along the length of the SACF and cannot be characterized simply in terms of a shift in the main peak in the SACF. When modeling the results of a number of psychophysical pitch studies, the location of the main peak in the SACF was used with some success as a predictor of pitch judgments (Meddis and Hewitt, 1991a). When modeling sensitivity to differences in phase relationships among harmonic components, comparisons along the length of the SACFs of the two stimuli were used (Meddis and Hewitt, 1991b). The latter procedure is a more general version of the former and probably a more secure approach. In the remaining evaluations of the model, comparisons will be made along substantial segments of the SACF.

B. Pitch discrimination

Houtsma and Smurzynski (1990) showed that the frequency difference limen for the pitch of an 11 component harmonic complex increased with the number of the lowest harmonic in the complex. Carlyon and Shackleton (1994) found that this is a function of the *resolvability* of constituent harmonics. We explored this issue by testing the model using Carlyon and Shackleton's (1994) paradigm. We used stimuli with harmonic complexes containing either resolved (LOW, from 125 to 625 Hz) or unresolved (HIGH, from 3900 to 5400 Hz) harmonics. Comparisons were made between tones with $F0$'s of 100 and 102 Hz. If the model is to simulate Carlyon and Shackleton's result we expect the 2% shift to produce a bigger change in the SACF when using low harmonics than when using high harmonics. To test this we calculated D^2 between the SACFs for the 100-Hz and the 102-Hz tones. The stimuli were generated using a signal consisting of 80 harmonics of the fundamental digitally filtered using a Butterworth filter set to give skirts of 24 dB per octave. The range was used to specify the 3 dB down points of the filter. The SACF was calculated after 10 complete cycles of the stimulus.

The SACFs for the LOW and HIGH harmonic stimuli are given in Fig. 5(a) and (c). The SACFs are very different in shape for the resolved (LOW) and unresolved (HIGH)
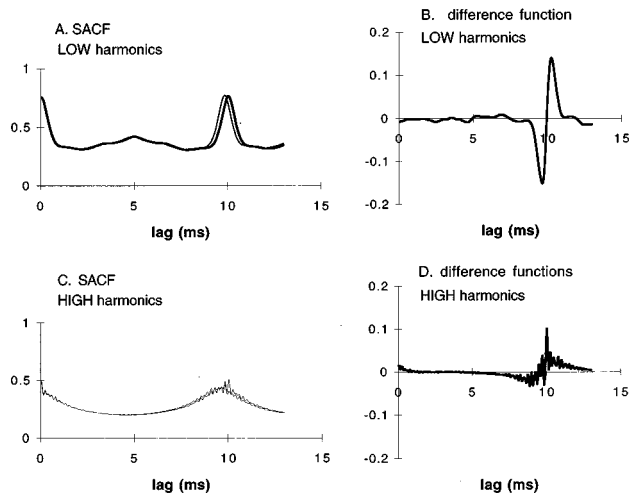
FIG. 5. SACFs and difference functions for signals with $F0$ of 100 and 102 Hz in the LOW and HIGH harmonic ranges. Top row, LOW harmonics; bottom row, HIGH harmonics. Left column, SACFs; right column, difference functions.

harmonics. The LOW stimulus has a narrow peak at a lag of 10 ms corresponding to a 100-Hz pitch percept. The HIGH stimulus has a broad and shallow peak at the same value.

A 2% shift in $F0$ creates very different effects for the two stimuli. For the LOW stimulus, a clear difference can be seen between the 100- and 102-Hz SACFs. For the HIGH stimulus, the difference is much less clear. Figure 5(b) and (d) shows the corresponding difference functions. The LOW stimulus produces a larger difference function when the pitch is shifted by 2%. Using the SACF range 80–1000 Hz, D^2 was more than three times as great for LOW (0.37) than for HIGH (0.11) harmonics confirming our expectations. D^2 was calculated here using a wide segment of the SACF (80–1000 Hz).

To demonstrate the generality and lawfulness of the effect, the above calculations were extended to include additional examples using a fundamental frequency of 200 Hz and two additional intermediate ranges of harmonics were studied. The four ranges are as follows: LOW (125–625 Hz), MID (1375–1875 Hz), MID/HIGH (2500–3500 Hz), and HIGH (3900–5400 Hz). Figure 6(a) shows the Euclidean distance for a 2% shift in $F0$ for all four harmonic ranges using two fundamental frequencies (100 and 250 Hz).

A large value of D^2 implies a high discriminability of the two signals and therefore our D^2 measure can be said to be loosely analogous to the d' statistic used in signal detection theory and employed by Carlyon and Shackleton. Our calculations based on the model are therefore qualitatively similar to their results (1994, Fig. 6) which show a deterioration in discriminability from LOW to HIGH ranges of harmonics.

Houtsma and Smurzynski (1990) measured frequency-difference limens using 11-component harmonic stimuli. They found an increase in threshold as the number of the lowest harmonic was raised from 7 to 19 for a 200-Hz fundamental. Their paradigm was simulated using an 80 harmonic stimulus band-pass filtered as described above; the lower and upper 3 dB down point of the filter were set to the
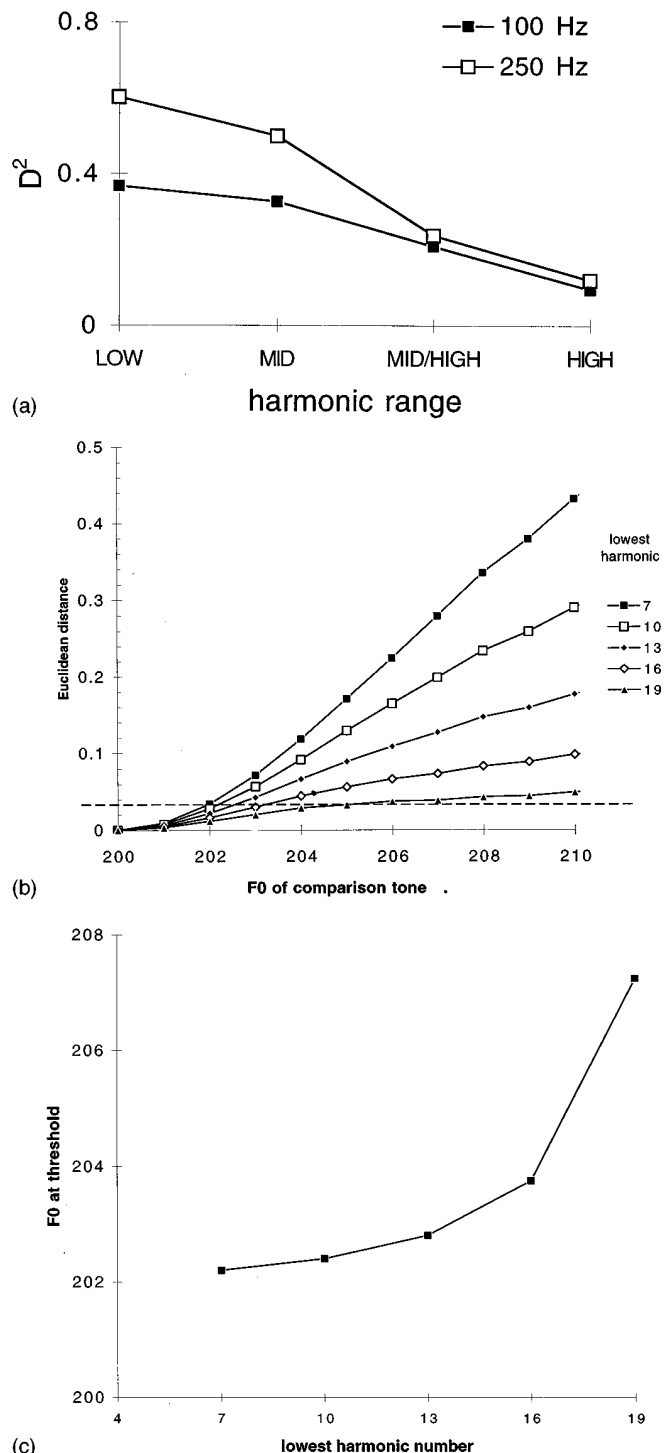
FIG. 6. (a) D^2 for pairs of stimuli with fundamental frequencies differing by 2%. Four different ranges of harmonics and two values of $F0$ are used (see text for details). (b) D^2 as a function of $\Delta F0$ and the number of the lowest harmonic. Tones consisted of 11 harmonics. The dashed horizontal line represents the criterion used to define "threshold." (c) "Threshold" values of $F0$ as a function of the number of the lowest harmonic. Values are derived from B using a criterion value of 0.042.

nth and the $(n+11)$th harmonic. Figure 6(b) shows D^2 as a function of $\Delta F0$ for different ranges of harmonics. Stimuli with low harmonic number show steeper slopes than those with high harmonic number. The Euclidean distance calcu-
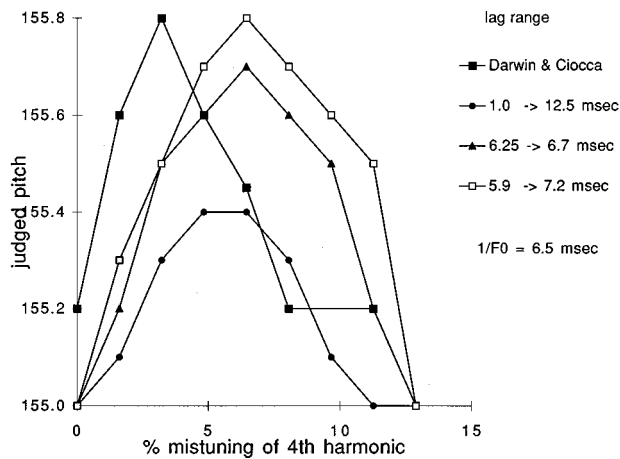
FIG. 7. Summary of model 'best match' decisions compared with Darwin and Ciocca's (1994) results (open circles). The model results are repeated with different lag ranges.

lations were based on a lag range corresponding to $F0$ $\pm 10\%$. When repeated using the wider range (80–1000 Hz), the pattern was similar.

The data in Fig. 6(b) can be represented more conventionally by setting a criterion value for D^2. A value of D^2 $= 0.042$ was chosen to give thresholds similar to listeners and the model thresholds are shown in Fig. 6(c). These show the expected upward trend but do not replicate the leveling out observed by Houtsma and Smurzinski for the highest range of harmonics. The four-stage model generates shallower, less distinctive SACFs for stimuli with unresolved components. As a result small changes in $F0$ have less effect on the SACF than when resolved harmonics are used.

C. The pitch of a complex tone with a mistuned harmonic

If one component of a harmonic complex is mistuned slightly, it gives rise to a shift in the pitch of the whole complex (e.g., Moore *et al.*, 1985). This effect is typically maximal for a 3%–4% mistuning of a low-numbered harmonic. The pitch shift can be as great as 1%. Darwin *et al.* (1994) have recently presented some highly detailed results for a harmonic stimulus with a fundamental frequency of 155 Hz. The behavior of the model is first illustrated using their paradigm.

To simulate Darwin *et al.*'s procedure, the SACFs were calculated for 10-component harmonic test tones of 100-ms duration with an $F0$ of 155 Hz and the fourth harmonic mistuned to a varying degree. To establish the best pitch match we also calculated the SACFs of a range of comparison 10-component harmonic tones, with $F0$ ranging between 155 and 156 Hz in steps of 0.1 Hz. For each test stimulus with a mistuned harmonic, the Euclidean distance was computed between its SACF and the SACFs of all comparison stimuli across the lag range corresponding to $F0 \pm 10\%$ (140–170 Hz). The comparison stimulus yielding the lowest value of D^2 was judged to be the best match.

Figure 7 summarizes these results by showing the frequency of the best pitch match as a function of the degree of mistuning of the fourth harmonic and compares it with data

from Darwin *et al.* (1994). Figure 7 also shows the results of repeating the same process using different lag ranges along the SACF. Changing the lag range causes quantitative changes but leaves the qualitative results intact. The biggest pitch shifts were obtained using a lag range of 5.9–7.1 ms ($1/F0 \pm 10\%$). This range will be used for the rest of this section. Darwin *et al.* obtained pitch matches of 155.2 Hz for no mistuning (when the fourth harmonic is 620 Hz). This represents a small bias of 0.2 Hz in the listeners' results. Allowing for this bias, the model results and the real data are in close agreement for the first three points.

When mistuning increases beyond 3%, the listeners' pitch matches decline in frequency while the model's pitch matches continue to rise in frequency. This may be attributable to an additional attentional mechanism beyond the scope of this study. As mistuning increases, listeners normally report that the mistuned harmonic can be heard as a separate entity. It is likely that this process of ''hearing out'' a mistuned component is accompanied by a weakening of the contribution of that component to the pitch percept (Darwin, 1992). We shall not pursue this mechanism further beyond noting that we have experimented elsewhere with a model for segregating sounds using pitch to guide the process (Meddis and Hewitt, 1992).

The return of the *model's* best match predictions to baseline levels when the mistuned harmonic is set to 700 Hz is a consequence of the fact that a mistuning of 700 Hz is both an upward mistuning of a fourth harmonic or an equal downward mistuning of a fifth harmonic; leading to a canceling of the effect.

D. The dominance of individual harmonics

Moore *et al.* (1985) explored the magnitude of the pitch shift following the mistuning of the first to the sixth harmonic. The results were variable across subjects but it was possible to conclude that mistuning the *second* harmonic by 3% had the greatest effect on the pitch of the complex for $F0$ of 100, 200, and 400 Hz [Fig. 8(b)]. The higher the harmonic (above the second) the less its influence. By contrast, the first harmonic was more influential at higher fundamental frequencies. We, therefore, extended the above evaluation to study the effect of mistuning the first to the sixth harmonics.

To avoid complications related to the ''hearing out'' phenomena discussed above, the evaluation is confined to 3% mistunings. The resulting pitch shifts are given in Fig. 8(a). The results were obtained using signal durations that were a complete number of cycles of the fundamental and close to 100 ms. D^2 calculations were based on lag ranges corresponding to $1/F0 \pm 10\%$.

While the observed shifts are quantitatively lower than Moore *et al.*'s results, the key characteristics of Moore *et al.*'s data are reproduced by the model. Mistuning the first harmonic has very little influence on the pitch of the complex with an $F0$ of 100 Hz but is increasingly influential at higher fundamental frequencies. The second harmonic is very influential for all three fundamental frequencies. Finally, the higher the harmonic (above the second) the less its influence on pitch.

Model

harmonic number

■ 1
□ 2
▦ 3
▨ 4
▥ 5
▨ 6

(a)

Moore et al.

harmonic number

■ 1
□ 2
▦ 3
▨ 4
▥ 5
▨ 6

(b)
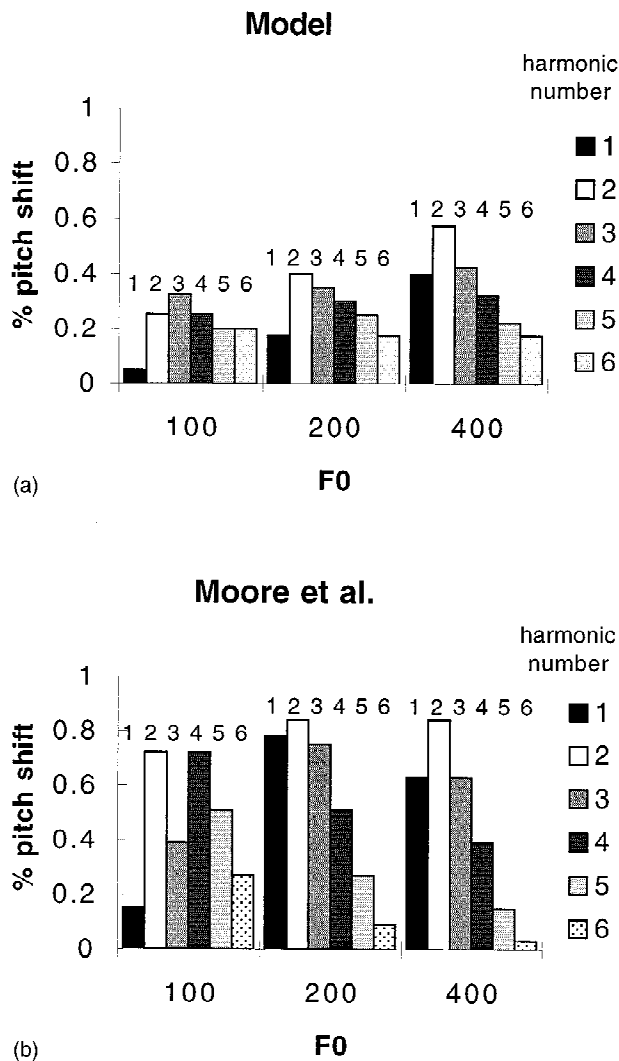
FIG. 8. Pitch shifts associated with 3% mistuning of harmonics 1–6 of 12-harmonic tones with $F0 = 100$, 200, and 400 Hz. (a) Model results; (b) Moore *et al.*'s (1985) results.

It is difficult to explain the discrepancy in size between the results of Moore *et al.* and the model output. The model's results are consistent with those of Darwin *et al.* (see above) who also produce smaller pitch shifts than Moore *et al.* for corresponding stimuli.

The very small pitch shifts found for the model when the first harmonic was mistuned are not untypical of the empirical results for an $F0$ of 100 Hz. We speculated that this might arise from the action of the model's pre-emphasis function which attenuates low-frequency components. However, when this was switched off, the effect remained intact.

Moore *et al.* (1985) explained their experimental results in terms of the principle of a ''dominance region.'' Ritsma's (1967, 1970) studies had earlier indicated that the third, fourth, and fifth harmonics made a dominant contribution to the pitch of a complex tone. Plomp (1967) suggested that the number of the dominant harmonic declined as the fundamental frequency increased. The model's behavior is consistent with Ritsma's dominance region except that we need to include the second harmonic in this region. Plomp's ideas are also represented in the model results; the third harmonic is

dominant when $F0$ is 100 Hz but the second harmonic is dominant at higher pitches. Similarly, the center of gravity of the three histograms moves toward lower harmonics as the pitch increases in both the model behavior and in the results of Moore *et al.*

The results in Fig. 8 were based on calculations using a lag range corresponding to frequencies within $\pm 10\%$ of $F0$. Additional calculations were made using a lag range corresponding to 80–1000 Hz. This resulted in some quantitative changes but the overall pattern of results and the conclusions that could be drawn were unchanged.

A careful examination of the model's operation leads to the suggestion that a number of principles interact to produce the dominance phenomenon. According to the model, the pitch sensation arises through the summation of the ACFs of the output of a bank of filters. At low frequencies each tone emerges from its corresponding filters as an unmodulated pure tone (at least for the fundamental frequencies studied here). The ACF of a pure tone can be represented as the cosine function. Low-frequency harmonics will have wider ACF peaks that will contribute more diffusely to the estimate of the most commonly occurring periodicity. According to this argument, higher-frequency harmonics should contribute more narrowly to the estimate and mistuning should disturb the pitch percept more. However, a second principle holds the first in check. As the frequency of the harmonic increases, the filtering which occurs at the level of the receptor potential reduces the ac component of its output. This reduces the amplitude of the cosine ACF and the extent of its contribution to the pitch percept. In Fig. 2 the peak-to-trough height of the ACFs can be seen to decrease in size with increasing filter center frequency. These two principles should produce an optimal harmonic number with reduced contributions from harmonics with higher and lower frequencies. This is certainly consistent with the results shown in Fig. 8. Because the channels are distributed on a logarithmic scale of frequency, high-frequency components are represented by a smaller number of channels and make a reduced contribution to the SACF. At the other end of the scale, the pre-emphasis of the outer and middle ear has the effect of reducing the amplitude of very low-frequency signals relative to middle-range frequencies above 1 kHz. While not critical to the principle of dominance, this latter effect will have some influence on the relative weights of stimulus components.

The principle of dominance, in the model at least, emerges from a set of no fewer than four subsidiary principles. It is doubtful whether the model is able to represent each effect accurately and, in any case, each principle will be subject to considerable individual differences in the real world.

III. DISCUSSION

These results extend the range of successful predictions made by the four-stage periodicity model. In particular, these results are consistent with the view, strongly supported by recent research, that the pitch percept based on resolved harmonics will behave differently from that based on unresolved harmonics. A principle emerges that resolved har-

monics exercise a potent influence on the pitch percept but their relative phases do not influence the percept. By contrast, unresolved harmonics make a weaker contribution to the percept but their relative phases are perceptually salient. The ability of our model to simulate qualitative aspects of these phenomena suggest that it may not be necessary to posit two distinct mechanisms of pitch extraction.

The four-stage model represents a general approach to pitch extraction. Our particular implementation is only one of a number of alternatives. The use of linear gammatone filters to simulate peripheral mechanical filtering and the hair cell model to represent the half-wave rectification and low-pass filtering are both open to improvement. Similarly, the use of autocorrelation for the periodicity stage of the model may seem an unlikely choice if the process is considered to represent some physiological process. It is likely, however, that many alternative algorithms sensitive to periodicity could be substituted here without prejudice to the operation of the model. The final summation stage is also open to further study; the method used here is the simplest we could devise. Finally, our method of pitch matching by using the Euclidean distance between an arbitrarily long segment of the SACF is undoubtedly open to further development as the model is challenged by new data. At present, it remains unclear which segment of the SACF should be used for predicting pitch judgments. It is the basic four-stage nature of the model which is critical to our explanation of pitch rather than any specific implementation.

The ability of the system to detect periodicity appears to be limited to relatively long periods. This means that the system cannot extract information from the *fine structure* of the input waveform in the high-frequency channels. However, usable periodicity information is still available in amplitude modulation of the envelope of the output of high-frequency filters. It is, therefore, a matter of expediency that the signal be reduced to its envelope by half-wave rectification and low-pass filtering (stage ii). The low frequency characteristics of the envelope can then be processed by the periodicity detection system. A similar phenomenon can be observed in relation to our sensitivity to interaural time differences (ITDs) in continuous signals. At low signal frequencies these can be detected in the ITDs of the fine structure but at high frequencies this sensitivity is restricted to ITDs in the *envelope* of the signal (Henning, 1974).

In these two respects, the model can be seen to come close to making sense of Carlyon and Shackleton's (1994) proposal that there are two distinct pitch extraction mechanisms. High-frequency, unresolved harmonic components interact with each other and are stripped of fine structure information so that pitch judgments must be made on the basis of the signal envelope. Low-frequency, resolved harmonics do not interact and have no useful envelope information. The HIGH and LOW sets of components give rise to distinctly different SACF shapes (see Fig. 4); the former contribute to a shallow diffuse peak while the latter support narrower and higher peaks. This observation is consistent with their assertion that frequency-difference limens are exceptionally high when the comparison stimulus consists of high harmonics and the test stimulus consists of low harmonics; in terms of

SACFs, we are not comparing like with like. It is doubtful, however, that this amounts to two distinct mechanisms.

Pitch matching and difference threshold phenomena have been simulated in this study using comparisons along whole segments of the SACF. An alternative approach is to measure the location of the highest peak and look for differences in location between the two stimuli. In early (unpublished) tests we found this method to require an unreasonably high degree of resolution for the SACF. However, in Fig. 5(b) it can be clearly seen that the large differences between the two SACFs occur to the left and right of the main peak. The sensitivity of our method depends on including these parts of the SACF in our difference measure.

Despite its predictive power, autocorrelation is unlikely to be the method used by the nervous system to detect periodicity. Physiological studies have, however, shown that a number of different cell types in the cochlear nucleus act as band-pass AM amplifiers (Moller, 1976; Frisina *et al.*, 1990a, b; Kim *et al.*, 1990; Rhode, 1995; Hewitt and Meddis, 1994). These units can also phase lock to low-frequency pure tones. These two properties make it likely that they may be involved in the periodicity detection process. Some cells in the central nucleus of the inferior colliculus show an enhanced rate of firing to restricted ranges of frequencies of amplitude modulation (Langner, 1992; Rees and Palmer, 1989). It is possible, but not yet proven, that each cell of this type corresponds to a single point in our autocorrelation function. While it remains to be established that the correspondence is valid, it is important to stress that autocorrelation is merely substituting *pro tem* for the periodicity detector until the most appropriate physiological mechanism can be determined.

Our proposal falls short of a general theory of pitch. It does not, for example, address the issues of pitch variation with sound-pressure level, octave enlargement, and interaural pitch differences. Nor have we attempted to accommodate the phenomena of binaural pitch perception (Houtsma and Goldstein, 1972; Beerends and Houtsma, 1989; Cramer and Huggins, 1958). Similarly, it has not been possible to address Hartmann's (1994a, b) suggestion that inharmonicity is detected locally among adjacent harmonics and cannot be detected if adjacent harmonics are missing.

ACKNOWLEDGMENTS

APPENDIX: PARAMETERS USED IN THE EVALUATION OF THE MODEL

Signal parameters
$1.00E-01$ Stimulus signal duration (seconds)—variable.
$5.00E-05$ Stimulus sampling interval, dt (seconds).
$2.500E-03$ Ramp up rise time for signal (seconds).

Pre-emphasis filter
2   Filter order.

| 450.0 | Lower, 3 dB down cutoff frequency (Hz). |
| 8500.0 | Upper, 3 dB down cutoff frequency (Hz). |

Gammatone filter parameters

4	Order for the gamma tone filters.
100	Lowest center frequency of gamma-tone filter bank.
8000	Highest center frequency of BM filter bank.
60	No. of filters.

Hair cell parameters

100.0	Permeability constant A (units per second).
6000.0	Permeability constant B (units per second).
2000.0	Release rate g (units per second).
5.05	Replenishment rate y (units per second).
2500.0	Loss rate l (units per second).
66.31	Reprocessing rate x (units per second).
6580.0	Recovery rate r (units per second).
1.0	Max. no. of transmitter packets in free pool m.
50 000.0	Firing rate scalar h (spikes per second)

Autocorrelation parameters

| $13.0e-3$ | maximum lag, l (seconds). |
| $10.0e-3$ | Time constant, τ (seconds). |

[1]LUTEar is a computational library produced to simplify and standardise the implementation of auditory simulation investigations. It is written in ANSI C in a modular format. The library contains a large selection of auditory models and analysis routines. It is available on application to the authors or directly from this WWW site: http://www.essex.ac.uk/psychology/hearinglab/

Assmann, P. F., and Summerfield, Q. (**1990**). ''Modeling the perception of concurrent vowels: vowels with different fundamental frequencies,'' J. Acoust. Soc. Am. **88**, 680–697.

Beerends, J. G., and Houtsma, A. J. M. (**1989**). ''Pitch identification of simultaneous diotic and dichotic two tone complexes,'' J. Acoust. Soc. Am. **85**, 813–819.

Brown, G. J., and Cooke, M. (**1994**). ''Perceptual grouping of musical sounds: A computational model,'' J. New Music Res. **23**, 1–25.

Carlyon, R. P., and Shackleton, T. M. (**1994**). ''Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms?,'' J. Acoust. Soc. Am. **95**, 3541–3554.

Cramer, E. M., and Huggins, W. H. (**1958**). ''Creation of pitch through binaural interaction,'' J. Acoust. Soc. Am. **30**, 413–417.

Darwin, C. J., Ciocca, V., and Sandell, G. J. (**1994**). ''Effects of frequency and amplitude modulation on the pitch of a complex tone and a mistuned harmonic,'' J. Acoust. Soc. Am. **95**, 2631–2636.

Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (**1990a**). ''Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement,'' Hearing Res. **44**, 99–122.

Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (**1990b**). ''Encoding of amplitude modulation in the Gerbil Cochlear Nucleus. II. Possible neural mechanisms,'' Hearing Res. **44**, 123–141.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ''Derivation of auditory filter shapes from notched-noise data,'' Hearing Res. **47**, 103–138.

Goldstein, J. L. (**1973**). ''An optimum processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. **54**, 1496–1517.

Hartmann, W. M. (**1994a**). ''Hearing a mistuned harmonic in an otherwise periodic complex tone,'' J. Acoust. Soc. Am. **88**, 1712–1724.

Hartmann, W. M. (**1994b**). ''On the perceptual segregation of steady-state tones,'' in *A Biological Framework for Speech Perception and Production*, edited by H. Kawahara (ATR Human Information Processing Laboratories, Kyoto, Japan).

Henning, G. B. (**1974**). ''Detectability of interaural delay in high-frequency complex waveforms,'' J. Acoust. Soc. Am. **55**, 84–90.

Hewitt, M. J., and Meddis, R. (**1994**). ''A computer model of amplitude modulation sensitivity of single units in the inferior colliculus,'' J. Acoust. Soc. Am. **95**, 2145–2159.

Houtsma, A. J. M., and Goldstein, J. L. (**1972**). ''The central origin of the pitch of complex tones: Evidence from musical interval recognition,'' J. Acoust. Soc. Am. **51**, 520–529.

Houtsma, A. J. M., and Smurzynski, J. (**1990**). ''Pitch identification and discrimination for complex tones with many harmonics,'' J. Acoust. Soc. Am. **87**, 304–310.

Kim, D. O., Sirianni, J. G., and Chang, S. O. (**1990**). ''Responses of DCN-PVCN neurons and auditory nerve fibers in unanesthetized decerebrate cats to AM and pure tones: Analysis with autocorrelation/power-spectrum,'' Hearing Res. **45**, 95–113.

Langner, G. (**1992**). ''Periodicity coding in the auditory system,'' Hearing Res. **60**, 115–142.

Licklider, J. C. R. (**1951**). ''A duplex theory of pitch perception,'' Experientia **7**, 128–133.

Licklider, J. C. R. (**1959**). ''Three auditory theories,'' in *Psychology: A Study of a Science*, edited by S. Koch (McGraw-Hill, New York).

Lyon, R. F. (**1984**). ''Computational models of neural auditory processing,'' IEEE Proc. 1984. 36.1.1–36.1.4.

Meddis, R. (**1986**). ''Simulation of mechanical to neural transduction in the auditory receptor,'' J. Acoust. Soc. Am. **79**, 702–711.

Meddis, R. (**1988**). ''Simulation of mechanical to neural transduction: Further studies,'' J. Acoust. Soc. Am. **83**, 1056–1063.

Meddis, R., and Hewitt, M. J. (**1991a**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification,'' J. Acoust. Soc. Am. **89**, 2866–2882.

Meddis, R., and Hewitt, M. J. (**1991b**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity,'' J. Acoust. Soc. Am. **89**, 2883–2894.

Meddis, R., and Hewitt, M. J. (**1992**). ''Modelling the identification of concurrent vowels with different fundamental frequencies,'' J. Acoust. Soc. Am. **91**, 233–245.

Meddis, R., Hewitt, M. J., and Shackleton, T. M. (**1990**). ''Implementation details of a computational model of the inner hair-cell/auditory-nerve synapse,'' J. Acoust. Soc. Am. **87**, 1813–1816.

Moller, A. R. (**1976**). ''Dynamic properties of primary auditory fibres compared with cells in the cochlear nucleus,'' Acta Physiol. Scand. **98**, 157–167.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985**). ''Relative dominance of individual partials in determining the pitch of complex tones,'' J. Acoust. Soc. Am. **77**, 1853–1860.

Moore, B. C. J. (**1977**). *An Introduction to the Psychology of Hearing* (Academic, New York).

Noorden, L. P. A. S. van (**1982**). ''Two channel pitch perception,'' in *Music, Mind and Brain*, edited by M. Clynes (Plenum, New York).

Patterson, R. D. (**1987**). ''A pulse ribbon model of monaural phase perception,'' J. Acoust. Soc. Am. **82**, 1560–1586.

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., Robinson, K. D., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (**1992**). ''Complex sounds and auditory images,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford).

Plomp, R. (**1967**). ''Pitch of complex tones,'' J. Acoust. Soc. Am. **41**, 1526–1533.

Rees, A., and Palmer, A. R. (**1989**). ''Neuronal responses to amplitude modulated and pure-tone stimuli in the guinea pig inferior colliculus and their modification by broadband noise,'' J. Acoust. Soc. Am. **85**, 1978–1994.

Rhode, W. S. (**1995**). ''Inter-spike intervals as a correlate of periodicity pitch in cat cochlear nucleus,'' J. Acoust. Soc. Am. **97**, 2414–2429.

Ritsma, R. J. (**1967**). ''Frequencies dominant in the perception of the pitch of complex sounds,'' J. Acoust. Soc. Am. **42**, 191–198.

Ritsma, R. J. (**1970**). ''Periodicity detection,'' in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden, The Netherlands).

Terhardt, E. (**1979**). ''Calculating virtual pitch,'' Hearing Res. **1**, 155–182

Schouten, J. F. (**1970**). ''The residue revisited,'' in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden).

Perceived lateral position of narrow-band noise in hearing-impaired and normal-hearing listeners under conditions of equal sensation level and sound-pressure level

Helen J. Simon and Inna Aleksandrovsky[a]
Smith–Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, California 94115

The perceived lateral position of narrow-band noise (NBN) was studied in a graphic pointer task as a function of the method of compensation for interaural threshold asymmetries in hearing-impaired and normal-hearing subjects. The method of compensation consisted of equal sensation level (EqSL) or equal sound-pressure level (EqSPL) at the two ears within the same subject. The NBN signals were presented at 11 center frequencies with interaural intensity differences (IIDs) that varied from -20 to $+20$ dB. When equalizing by SL, the perceived lateral position is essentially linearly dependent on the degree and direction of asymmetry in asymmetric normal-hearing and hearing-impaired listeners. Equalizing by SPL shows no such dependency but produces images that are lateralized close to the midline. These results reveal that subjects may have adapted to their threshold asymmetries. These results will be discussed in terms of the fitting of binaural hearing aids. © *1997 Acoustical Society of America.* [S0001-4966(97)03809-5]

PACS numbers: 43.66.Pn, 43.66.Qp [RHD]

INTRODUCTION

The purpose of this paper is to report the results of a study that systematically investigated the relation between lateralization performance in subjects with normal hearing and sensorineural hearing loss (SNHL) when signals with interaural intensity differences (IIDs) were presented with equal sensation level (EqSL) or equal sound-pressure level (EqSPL) at the two ears. Previous work with normal-hearing listeners that assumed left–right symmetry (Hershkowitz and Durlach, 1969; Domnitz, 1973) showed reduced performance in certain binaural discrimination and detection tasks with imbalanced as compared to balanced intensity input. With the notable exception of Koehnke *et al.* (1995), there has not been compensation for threshold asymmetries in binaural experiments with SNHL listeners (Durlach *et al.*, 1981; Colburn *et al.*, 1987). It is assumed that degradation in binaural performance in asymmetric subjects might be due to loudness or SL imbalances as occurs with input imbalances in normals. Durlach *et al.* (1981) suggested that some method of equating interaural stimulus amplitude might be needed to restore the binaural performance and discussed four methods: (a) EqSPL, (b) EqSL, (c) equal loudness (EqL) at the two ears, and (d) a centered image. They argued that, for normal-hearing listeners (and we assume for symmetric thresholds), all four methods would ''roughly'' equal 0 dB IID. However, for asymmetric SNHL, the interaural stimulus amplitude for EqSL and EqSPL would differ from each other. Further, EqSPL would be the closest to and EqSL would be the furthest from a centered image due to, in the former, the correlation of auditory, visual, and tactile perception and sensorimotor feedback (Durlach *et al.*, 1981).

Hawkins and Wightman (1980) and Smoski and Trahiotis (1986) utilized EqSPL and EqSL with normal-hearing and SNHL listeners. Generally, interaural time delay (ITD) thresholds for narrow-band noise (NBN) centered on 500 Hz revealed better performance in the EqSPL condition than in the EqSL condition, with SNHL listeners only slightly less sensitive than normal-hearing listeners. Thresholds were higher for both groups for the 4000-Hz signals than for the 500-Hz signals. At 4000 Hz, normal-hearing listeners again showed superior performance in the EqSPL condition while performance for the SNHL subjects was approximately equal in the EqSL and EqSPL conditions.

Smoski and Trahiotis (1986) noted that the degradation of ITD thresholds in the high frequencies at EqSL may be the result of the change in the overall intensity level of the stimulus. Especially for the normal-hearing listeners, the intensity level varied as a function of EqSPL and EqSL: EqSPL was presented at a higher intensity level (always 80 dB SPL) than EqSL (25 dB SL).

The present study investigates the perceived lateral position of NBN as a function of balancing the inputs to both ears by EqSPL or EqSL in normal-hearing and SNHL subjects. Subjects were tested at 25 dB SL (EqSL) *re:* NBN threshold. However, with threshold asymmetries between the two ears, the results at EqSL would be unbalanced or unequal (UnEq) in SPL. Thus at those frequencies where inputs presented in the EqSL condition were asymmetrical in the two ears (UnEqSPL), the subjects were retested with ''balanced'' SPL inputs achieved by bilaterally presenting an intensity level corresponding to 25 dB SL at the poorer ear.[1] This balanced condition resulted in EqSPL (UnEqSL).[2] The decision to balance in this manner was made to insure that the signal in the poorer ear was at 25 dB SL, thus allowing a direct comparison with the EqSL condition.

Thus this study extends current knowledge in that the intensity levels for EqSL and EqSPL are consistent within the same subject, interaural asymmetries are documented and tested, and the interaural asymmetries in the EqSPL condition are ''balanced.''

---

[a]Now at Dept. of Psychology, University of California, Berkeley, California.

TABLE I. Narrow-band noise (NBN) presentation levels (dB SPL) at the 11 center frequencies for the left/right ears. These levels were obtained by adding 25 dB to the NBN thresholds. An asterisk indicates those frequencies where hearing thresholds are not within normal limits for pure tones (*re:* ANSI, 1969). A double asterisk indicates those frequencies not tested experimentally since the subject could not tolerate the EqSPL condition.

Center frequency (Hz)	Normal		Hearing impaired				
	LP	LT	SA	AB	JD	EG	LAH
250	40/45	35/40	40/50	45/40	60/55	65/75*	50/50
300	45/40	35/40	40/50	45/40	55/55	70/80*	45/50
400	45/40	40/35	40/45	35/35	50/55	70/75*	50/50
500	45/50	35/35	35/35	35/35	50/50	70/75*	45/40
750	40/40	30/25	30/35	30/35	45/40	70/70*	45/45
1000	25/35	35/30	25/35	40/35	40/50	65/70*	50/45
1250	25/25	25/35	25/25	35/30	45/45	65/70*	75/65*
1500	30/25	30/30	30/25	40/40	50/45*	70/70*	100/85*
2000	35/35	35/40	40/40	65/55*	60/60*	80/80*	90/90*
3000	30/25	25/25	60/75* **	90/85*	75/80*	75/80*	85/75*
4000	25/30	45/30	80/90* **	90/90*	85/90*	75/75*	85/75*

I. PROCEDURE

A. Subjects

Three female and two male subjects with SNHL and one female and one male with normal hearing participated in this study. Audiometric evaluations, using the Method of Limits, were performed using pure tones and NBN signals. Table I shows the 25 dB SL level (in dB SPL) for the NBN signals. As shown in this table, asymmetries for both groups of subjects were between 5 and 15 dB. In order to be insensitive to daily fluctuations, thresholds were determined in 5-dB steps and NBN thresholds were established immediately prior to the experiment. Six of the subjects had previous experience in psychoacoustic experiments in this laboratory and in addition, all subjects underwent extensive practice using this paradigm. The training was completed when at least one stimulus sequence at all frequencies was completed and performance was considered stable (usually 6–12 sessions).

B. Stimuli

Using the ''quadrature'' noise method described by Amenta *et al.* (1987), 100-Hz bandwidth NBN Gaussian pseudorandom noise stimuli were digitally generated offline. One bandwidth value was used in order to keep the time envelope fluctuations constant at different frequencies, even though the resulting bandwidths differ in critical band units.

Stimuli consisted of NBNs with center frequencies of 0.25, 0.3, 0.4, 0.5, 0.75, 1, 1.25, 1.5, 2, 3, and 4 kHz. Ten tokens at each center frequency were created. Each token was a 100-ms sample of an NBN including 20-ms on/off raised cosine function ramps. Tokens were digital-to-analog converted at a 32 000-Hz sampling rate with 16 bits of resolution using two separate channels. Both the overall presentation level and the magnitude of the IID were controlled as needed in 0.3-dB steps with a digitally controlled two-channel attenuator. The IIDs were created symmetrically by increasing one channel and decreasing the other. After digitization, the signal was low-pass filtered using a Chebyshev fourth-order reconstruction filter, with 10.5 kHz as the corner frequency, that yielded 90-dB attenuation at 16 kHz.

C. Methods

Signals were presented binaurally through ER 3-A insert earphones[3] to individual subjects seated in an IAC sound-attenuated room. Subjects were tested at 25 dB SL *re:* NBN threshold at EqSL and EqSPL.

If the stimuli were reported as diffuse (e.g., Domnitz and Colburn, 1977), subjects were instructed to locate the perceived center of the intracranial image associated with the stimulus. Each one and one-half hour session began with a minimum of six practice trials preceding the experimental runs.

D. Experimental paradigm

The experimental paradigm was a modification of a psychoacoustical mapping method such as that used by Blauert and Lindemann (1986). The IIDs of the stimuli were varied randomly in 4-dB steps over a range of ±20 dB at each of 11 frequencies. Each trial consisted of three 100-ms stimuli with 50-ms interstimulus intervals. Each triplet was followed by a 400-ms pause. These triplets, repeated until the subject responded, constituted one trial. The subjects indicated where they heard the stimulus by pointing a mouse-controlled cursor to the position on a schematic of the front of a head depicted on the computer screen. The computer program converted the mouse position units to a range of 200 arbitrary units, with 0 calibrated as midline and ±100 at the two ears.

An experimental run consisted of eleven trials, one trial at each of the 11 IIDs at one frequency. Center frequency and condition of balance were held constant over an experimental run. Each subject completed a minimum of ten runs. Each condition took several sessions to complete, thus insuring that data for each frequency and balance condition were gathered on more than one day.

II. RESULTS

A repeated measures analysis of variance (ANOVA) was carried out to examine the effects of subjects, group, center frequency, balance, degree of asymmetry, better ear,

TABLE II. Repeated measure analysis of variance for EqSPL versus EqSL data. Only the results of the main effects and significant three-way interactions pertaining to the balancing conditions are shown. An asterisk indicates the results significant at a criterion level of $p<0.05$.

Source of variance	Degrees of freedom	Sum of squares	F ratio
Subject	5	18 847.98	10.48*
Group (hearing impaired and normal)	1	1358.13	0.36
Balance	1	18.23	0.05
Degree of asymmetry	1	1050.85	2.92
Center frequency	1	926.83	2.58
IID	1	9608.34	26.71*
Better ear	1	741.80	2.06
Degree of asymmetry*balance*better ear	1	4526.82	12.59*
IID*degree of asymmetry*balance	1	1761.89	4.90*

FIG. 1. The effect of IID and degree of asymmetry on perceived intracranial position in EqSL and EqSPL. The ordinate represents the perceived lateral position of the IID stimulus, with 0 calibrated as perceived midline and ± 100 perceived at the two ears. Positive values are to the right of midline and negative values to the left. IID in dB is shown on the abscissa with 0 indicating equal intensive stimuli, positive values indicating stimuli more intense in the right ear, and negative values indicating stimuli more intense in the left ear. The parameter is the degree of asymmetry in dB SPL [panel (a)] or dB SL [panel (b)]. Each point represents the mean for the seven subjects. Error bars indicate the standard deviation of the mean.

and IID on perceived lateral position. Individual subject data points that were more than three standard deviations from the group mean at each IID were considered outliers and were removed for statistical calculation. A type III sum-of-squares was used since the number of observations behind each data point was not always identical. It can be noted that not all subjects had asymmetries or unbalanced data at each frequency. A criterion level of $p<0.05$ is used throughout to assess statistical significance. The results of the main effects and significant three-way interactions pertaining to the balancing conditions are shown in Table II.

The only significant main effect is IID indicating that for larger IIDs the position of the signal is perceived farther away from the center of the head (see Fig. 1). There were other significant two- and three-way interactions observed, however, only the interactions pertaining to the balancing conditions will be discussed since the other interactions do not affect the significant effects detected with respect to balance.

Figure 1 illustrates the significant three-way interaction $[F (1,360)=4.90; p=0.0271]$ as a function of the degree of asymmetry, IID, and balance. The ordinate represents the perceived mean position of the intracranial image, with 0 calibrated as perceived midline and ± 100 perceived at the two ears. Positive values are to the right of midline and negative values to the left. IID in dB is shown on the abscissa with 0 indicating equal intensive stimuli, positive values indicating stimuli more intense in the right ear, and negative values indicating stimuli more intense in the left ear. The parameter is the degree of asymmetry in the two balance conditions, dB SPL [panel (a)] or dB SL [panel (b)]. Each point represents the mean for the seven subjects. There was no significant difference between the two groups of subjects when a group was added in a four-way interaction and thus the data were combined in this and the following figure. Error bars indicate the standard deviation of the mean. Threshold differences favoring right or left ears were combined for this graph, i.e., each point combined the conditions where right ear input was 5 (10, 15) dB greater or 5 (10, 15) dB less than the left ear input. The resultant mean position of the intracranial image was shifted towards the ear where more subjects had poorer thresholds.

This figure shows that in general, for larger IIDs, the position of the signal is perceived farther away from the center of the head. In the EqSPL condition, the perceived lateralization is relatively unaffected by the degree of asymmetry. However, in the EqSL condition, the dependency on the degree of asymmetry is more notable, especially with the 10- and 15-dB asymmetries.

Figure 2 illustrates the perceived mean position of the intracranial image as a function of the significant three-way interaction of degree of asymmetry, better ear, and balance for the EqSPL and EqSL conditions $[F (1,360)=12.59; p=0.0004]$. Note that although this effect is represented as a three-way interaction in the ANOVA, the figure combines ''better ear'' and ''degree of asymmetry'' into one signed variable for salience of interpretation. The interaural input asymmetry is shown on the abscissa, with 0 indicating no
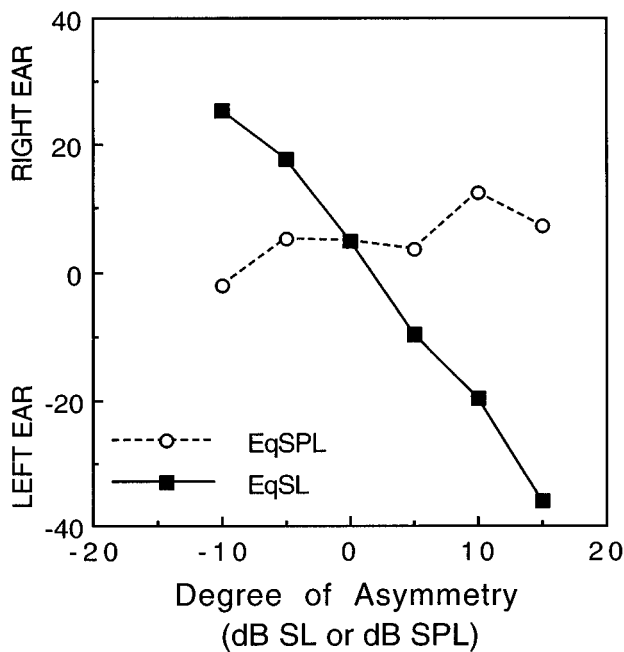
FIG. 2. Mean position of the intracranial image as a function of degree of asymmetry in SL and SPL. The vertical axis indicates the mean perceived horizontal (lateral) position of the intracranial position; the horizontal axis indicates the degree of interaural input asymmetry with 0 indicating no difference between the ears and positive and negative values indicating greater intensity (consistent with the poorer ear) to the left and right ears, respectively. The parameters are conditions EqSPL and EqSL.

difference between the ears and positive and negative values indicating greater intensity (consistent with the poorer ear) to the left and right ears, respectively. The ordinate indicates the mean perceived horizontal (lateral) position of the intracranial image. Positive and negative values indicate percepts to the right and left of midline, respectively. The parameters are conditions EqSPL and EqSL. Each data point represents one of the six values of asymmetry (± 15, ± 10, ± 5 dB), the mean of at least 1500 repetitions for the normals and a minimum of 264 (maximum 1243) for the SNHL.[4] Figure 2 demonstrates that for EqSL, lateral position is essentially linearly dependent on the degree of asymmetry. For EqSPL, again there is essentially no dependency of lateralization position on the degree of asymmetry.

III. DISCUSSION

This study confirms the assumptions of Durlach et al. (1981) for SNHL listeners and the results of Simon et al. (1994) for normal-hearing listeners: EqSPL is closest to, and EqSL is the furthest from a centered image. In addition, this result appears to be independent of the degree of asymmetry (to ± 15 dB) for EqSPL and is similar for normal and SNHL listeners.

These findings are also consistent with those of Hawkins and Wightman (1980) and Smoski and Trahiotis (1986)[5] where subjects generally show superior performance at EqSPL. The present results augment a previous report from this laboratory. In an acoustic pointing task (Simon et al., 1994), when slight asymmetries at the two ears produced signals of EqSL (UnEqSPL), lateralization was towards the

ear with the greater SPL signal, regardless of the ear to which the signal was delayed by ITD. The result with the acoustic pointing task and the present study suggest that people with asymmetric normal hearing have adapted to their asymmetry for IID and ITD lateralization tasks at individual frequencies.

The role of adaptation in lateralization experiments whether positioned by ITD or IID has also been documented by Bauer et al. (1966) and Florentine (1976). In studies of real and simulated asymmetrical hearing loss, Bauer et al. (1966) and Florentine (1976), respectively, found that at the onset of the hearing loss (or occlusion) subjects used EqL to center the image. After a period of time, the subjects in the former study used EqSPL to center the image while subjects in the latter study used a setting somewhere between EqSPL and EqL at the two ears. Nabelek et al. (1980) and Byrne and Dermody (1975) tested subjects with balanced (EqSPL) and unbalanced (UnEqSPL) gains in hearing aids. Both found that to produce a centered image, some of their subjects ignored the imbalance between the aids, an example of adaptation in hearing aid fitting.

An important issue as a result of these studies and the present findings concerns the current practices of fitting ''binaural'' hearing aids. The common assumption that two optimally fit monaural hearing aids constitute an optimum binaural fit is not necessarily true (Braida et al., 1979; Colburn et al., 1987; Koehnke and Zurek, 1990). The importance of the true ''binaural,'' and insufficiency of simply ''bilateral'' hearing aid fitting has been recently emphasized (Sandlin, 1994). Some interaction between both hearing aids is necessary to access the binaural advantage. However, what constitutes this interaction is unclear. Kimberly et al. (1994) suggest that adjustments of the interaural amplitude ratio are necessary to compensate for an asymmetrical loss. Schweitzer (1993) and Jerger et al. (1993) consider the phase relationships between the two hearing aids to be important and suggest variable phase adjusters for varying phase relationships to reflect individual differences. Thus some relationship between the two hearing aids in either time or intensity is being advocated.

However, no prescriptive procedures to date specifically document the appropriate method of adjusting the interaural levels in binaural hearing aids. In the traditional hearing aid fitting, the amount of real-ear gain prescribed by the prescription formulas (the actual gain provided by the hearing aid for an individual) in specific frequency regions is determined by data from one of two sources, threshold or comfort. If threshold procedures are used (i.e., Byrne and Cotton, 1988; McCandless and Lyregaard, 1983) asymmetries between the ears may result in unequal gain prescribed to each ear. If most comfortable loudness (MCL) procedures are used, even bilaterally symmetrical thresholds would not necessarily produce equal MCL in the two ears (cf. Lee, 1995) and in those instances the formulas also would prescribe bilaterally asymmetrical gain. To the extent that lateralization performance can predict localization abilities, the resultant SPL imbalance using these prescription formulas would disrupt the previously adapted system and impair localization and binaural release from masking. Presently in clinical prac-

tice, when threshold asymmetries exist, there is intuitive knowledge that ''asymmetrical'' gain fittings will cause perceptual ''imbalance.'' Therefore attempts are made to ''underamplify'' the worst ear to bring the gain functions for both ears as close to each other as the hearing loss permits. At the same time, the patient is often asked to ''balance'' the percept of the sound source or to adjust ''user gain'' by manipulating the volume controls on both hearing aids.[6] To our knowledge, there has not been a systematic study of the adjusted outputs in terms of the user gain relationship of binaural hearing aids. A possibility exists that a patient will, in effect, try to counterbalance the prescription-induced interaural SPL differences, achieving some UnEqSL, minimally UnEqSPL listening conditions to which he or she is adapted. The procedure involves balancing the loudness percept of a broadband signal like speech, leaving the interaural intensity/loudness relationships of the individual frequencies unaccounted for and possibly disrupted.

While the bilateral hearing aid user is likely to adapt to these prescription-induced imbalances for subjective tasks like perceptual centering, Colburn *et al.* (1987) note that long-term adaptation to the imbalance of the inputs does not occur for sensitivity tasks. By applying different gain functions to asymmetrical ears, a listener with SNHL could be at a perceptual disadvantage, whether temporary or permanent. Although listeners report better speech perception and localization ability with binaural as compared to monaural amplification (Koehnke *et al.*, 1992), the less-than-optimal improvement and acceptance of binaural amplification in some cases may be the result of these imbalances.

In conclusion, balancing by SPL rather than SL produced more accurate lateralization performance in asymmetric normal-hearing and SNHL listeners. Thus in binaural experiments, compensation for threshold asymmetries by equalizing SL between the two ears with SNHL listeners may not be necessary. However, our results do indicate the importance of testing thresholds in normal-hearing subjects and balancing for dB SPL in binaural experiments. Clinically, more attention and research is needed on the binaural rather than the bilateral fitting of hearing aids.

[1]The decision to balance the ears for EqSPL by presenting an intensity level corresponding to 25 dB SL at the poorer ear was done so that both ears were at least at 25 dB SL. Presenting an intensity level corresponding to 25 dB SL in the better ear would reduce the sensation level in the poorer ear below 25 dB. The manner of matching the ears will determine which signals are presented at the greater overall level.

[2]EqSPL or balance could also occur naturally when the thresholds in the two ears were equal at a particular frequency (EqSPL and EqSL). These differences will be discussed in relation to a different data set and paradigm in subsequent publications.

[3]These earphones were chosen because of their good intrasubject reliability in comparison to supra-aural phones: they reduce the variability of the volume of trapped air between the transducer and the eardrum (Wilber *et al.*, 1988; Killion and Villchur, 1989) and eliminate variations caused by earphone placement or collapsed ear canals, a common problem in the elderly (Olsho *et al.*, 1985). The insert earphones were positioned only once during a session. The insert earphones were inserted 2 to 3 mm past the canal entrance (a deep insertion), and visually checked by the investigator. Little test variability between sessions was anticipated from earphone placement (Wilber *et al.*, 1988).

[4]Since mean position of the intracranial image for the full-range IID yields polar opposite results, variability statistics for the collapsed IID data are not meaningful.

[5]Hawkins and Wightman (1980) tested three normal-hearing subjects with reported asymmetries of 10 dB or less. At 500 Hz, three of their eight SNHL subjects had 5-dB asymmetries and two had asymmetries of 40 dB or greater. At 4000 Hz, four subjects had asymmetries of 5 dB and three others had 10-, 35-, and 45-dB asymmetries. Smoski and Trahiotis (1986) did not report on the asymmetries of their two normal-hearing subjects. However, three of the four hearing-impaired subjects tested had asymmetries of 5 and 10 dB at 500 Hz and 5 and 25 dB at 4000 Hz.

[6]The validity of such a procedure is seldom discussed in the literature. However, this clinical judgment is validated by our results.

Amenta, C. A., Trahiotis, C., Bernstein, L. R., and Nuetzel, J. M. (**1987**). ''Some physical and psychological effects produced by selective delays of the envelope of narrow bands of noise,'' Hearing Res. **29**, 147–161.

ANSI (**1989**). ANSI S3.6-1989, ''Specifications for audiometers'' (American National Standards Institute, New York).

Bauer, R. W., Matuza, J. C., Blackmer, R. F., and Glucksberg, S. (**1966**). ''Noise lateralization after unilateral attenuation,'' J. Acoust. Soc. Am. **40**, 441–444.

Blauert, J., and Lindermann, W. (**1986**). ''Spatial mapping of intracranial auditory events for various degrees of interaural coherence,'' J. Acoust. Soc. Am. **79**, 806–813.

Braida, L. D., Durlach, N. I., Lippmann, R. P., Hicks, B. L., Rabinowitz, W. M., and Reed, C. M. (**1979**). *Hearing Aids-A Review of Past Research on Linear Amplification, Amplitude Compression and Frequency Lowering* (ASHA, Rockville, MD).

Byrne, D., and Cotton, S. (**1988**). ''Evaluation of the National Acoustic Laboratories' new hearing aid selection procedure,'' J. Speech Hear. Res. **31**, 178–186.

Byrne, D., and Dermody, P. (**1975**). ''Localization of sound with binaural body-worn hearing aids,'' Br. J. Audiol. **9**, 107–115.

Colburn, H. S., Zurek, P. M., and Durlach, N. I. (**1987**). ''Binaural Directional Hearing—Impairments and Aids,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York), pp. 261–278.

Domnitz, R. H. (**1973**). ''The interaural time JND as a simultaneous function of interaural time and interaural amplitude,'' J. Acoust. Soc. Am. **53**, 1549–1552.

Domnitz, R. H., and Colburn, H. S. (**1977**). ''Lateral position and interaural discrimination,'' J. Acoust. Soc. Am. **61**, 1586–1598.

Durlach, N. I., Thompson, C. L., and Colburn, H. S. (**1981**). ''Binaural interaction in impaired listeners,'' Audiol. **20**, 181–211.

Florentine, M. (**1976**). ''Relation between lateralization and loudness in asymmetrical hearing losses,'' J. Am. Aud. Soc. **1**, 243–251.

Hawkins, D. B., and Wightman, F. L. (**1980**). ''Interaural time discrimination ability of listeners with sensorineural hearing loss,'' Audiol. **19**, 495–507.

Hershkowitz, R. M., and Durlach, N. I. (**1969**). ''Interaural time and amplitude JNDS for a 500-Hz tone,'' J. Acoust. Soc. Am. **46**, 1464–1467.

Jerger, J., Lew, H. L., Chmiel, R., and Silman, S. (**1993**). ''Letter to the Editor: Confounding binaural interactions. Response to Schweitzer,'' J. Am. Acad. Audiol. **4**, 276.

Killion, M., and Villchur, E. (**1989**). ''Comments on 'Earphones in audiometry','' J. Acoust. Soc. Am. **85**, 1775–1778.

Kimberly, B. P., Dymond, R., and Gamer, A. (**1994**). ''Binaural digital hearing aids for binaural hearing,'' ENT J. **73**, 176–179.

Koehnke, J., and Zurek, P. M. (**1990**). ''Localization and binaural detection with monaural and binaural amplification,'' J. Acoust. Soc. Am. Suppl. 1 **88**, S169.

Koehnke, J., Besing, J., and Goulet, C. (**1992**). ''Speech intelligibility, localization, and binaural detection with monaural and binaural amplification,'' J. Acoust. Soc. Am. **92**, 2434(A).

Koehnke, J., Culotta, C. P., Hawley, M. L., and Colburn, H. S. (**1995**). ''Effects of reference interaural time and intensity differences on binaural performance in listeners with normal and impaired hearing,'' Ear Hear. **16**, 331–353.

Lee, L. W. (**1995**). Are the most-comfortable-loudness level measures the same between the ears for individuals with symmetrical sensorineural hearing loss? In *Hearing Aid Research and Development* (NIDCD and Dept. Vet. Affairs, Bethesda, MD), p. 42.

McCandless, G. A., and Lyregaard, P. E. (**1983**). ''Prescription of gain/output (POGO) for hearing aids,'' Hear. Instrum. **35**, 16–21.

Nabelek, A. K., Letowski, T., and Mason, D. (**1980**). ''An influence of binaural hearing aids on positioning of sound images,'' J. Speech Hear. Res. **23**, 670–687.

Olsho, L. W., Harkins, S. W., and Lenhardt, M. L. (**1985**). ''Aging and the auditory system,'' in *Handbook of the Psychology of Aging*, edited by J. Birren and K. W. Schiver (Van Nostrand Reinhold, New York), pp. 332–377.

Sandlin, R. E. (**1994**). ''Fitting binaural amplification to asymmetrical hearing loss,'' in *Strategies for Selecting and Verifying Hearing Aid Fittings*, edited by M. Valente (Thieme, New York), pp. 207–227.

Schweitzer, C. (**1993**). ''Confounding binaural interactions,'' J. Am. Acad. Audiol. **4**, 275–276.

Simon, H. J., Collins, C. C., Jampolsky, A., Morledge, D. E., and Yu, J. (**1994**). ''The measurement of the lateralization of narrow bands of noise: The effect of sound-pressure level,'' J. Acoust. Soc. Am. **95**, 1534–1547.

Smoski, W., and Trahiotis, C. (**1986**). ''Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss,'' J. Acoust. Soc. Am. **79**, 1541–1547.

Wilber, L. A., Kruger, B., and Killion, M. C. (**1988**). ''Reference thresholds for the ER-3A insert earphone,'' J. Acoust. Soc. Am. **83**, 669–676.

Speech intelligibility enhancement using hearing-aid array processing

Gabrielle H. Saunders[a] and James M. Kates[b]
Center for Research in Speech and Hearing Sciences, City University of New York, Graduate Center, Room 901, 33 West 42nd Street, New York, New York 10036

Microphone arrays can improve speech recognition in noise for hearing-impaired listeners by suppressing interference coming from other than the desired signal direction. In a previous paper [J. M. Kates and M. R. Weiss, J. Acoust. Soc. Am. **99**, 3138–3148 (1996)], several array-processing techniques were evaluated in two rooms using the AI-weighted array gain as the performance metric. The array consisted of five omnidirectional microphones having uniform 2.5-cm spacing, oriented in the endfire direction. In this paper, the speech intelligibility for two of the array processing techniques, delay-and-sum beamforming and superdirective processing, is evaluated for a group of hearing-impaired subjects. Speech intelligibility was measured using the speech reception threshold (SRT) for spondees and speech intelligibility rating (SIR) for sentence materials. The array performance is compared with that for a single omnidirectional microphone and a single directional microphone having a cardioid response pattern. The SRT and SIR results show that the superdirective array processing was the most effective, followed by the cardioid microphone, the array using delay-and-sum beamforming, and the single omnidirectional microphone. The relative processing ratings do not appear to be strongly affected by the size of the room, and the SRT values determined using isolated spondees are similar to the SIR values produced from continuous discourse. © *1997 Acoustical Society of America.* [S0001-4966(97)00709-1]

PACS numbers: 43.66.Ts, 43.60.Gk [JWH]

INTRODUCTION

In this paper, delay-and-sum beamforming and superdirective array processing are evaluated for a short microphone array suitable for hearing-aid applications. The primary objective of the work is to test the array effectiveness in improving speech intelligibility under typical listening situations. A secondary objective is to compare the observed benefit with predictions made from physical array measurements. The array processing is compared to a single omnidirectional microphone and a single directional microphone having a cardioid spatial response pattern. The processing effectiveness is evaluated using the speech reception threshold (SRT) and speech intelligibility rating (SIR) for a group of hearing-impaired subjects. The SRT uses a closed set of isolated spondees, while the SIR uses continuous discourse. An office and conference room were used for the data acquisition. The array was mounted in an endfire orientation on a microphone stand. The data from real rooms avoids the limitations of computer simulations that are often used in evaluating array-processing algorithms, and permits an accurate comparison of different microphone and array designs that have been proposed for hearing aids.

Evaluations of high-resolution and adaptive hearing-aid microphone arrays have usually been carried out using acoustic measurements of the array performance (Greenberg

and Zurek, 1992; Kates and Weiss, 1996). Subject data from such evaluations using speech intelligibility tests under realistic listening conditions are limited (Soede *et al.*, 1993a,b; Hoffman *et al.*, 1994; Kompis and Dillier, 1994). Despite the fact that test conditions and processing strategies have varied, the intelligibility studies have consistently shown that microphone arrays improve speech intelligibility for both normal hearing and hearing impaired subjects. Soede *et al.* (1993a) made acoustic measurements of a microphone array that was 10 cm long and consisted of five cardioid microphones, the outputs of which were combined using delay-and-sum beamforming. A diffuse noise field was generated by eight loudspeakers placed in the corners of a test booth. They then tested normal-hearing and hearing-impaired subjects with sentence materials using this array (Soede *et al.*, 1993b). Their acoustic measurements showed an improvement in signal-to-noise ratio (SNR) of between 5–12 dB, with the greatest improvement occurring at the highest frequencies, while subjects' SRTs improved by an average of 7 dB. The array used by Soede *et al.* can be hand-held or built into an eyeglass frame, and the array performance does not appear to be affected to any great extent by being mounted on the head.

Hoffman *et al.* (1994) tested normal-hearing subjects using spondees as stimuli and a seven-microphone adaptive array positioned around the head. They used a computer simulation to provide different degrees of room reverberation for a single source of interference at 45° to the array axis. Their acoustic measurements showed improvements in intelligibility-weighted SNR ranging from 22.7 dB under

---

[a]Present address: Scientific Learning Corporation 417 Montgomery St., Suite 500, San Francisco, CA 94104. Electronic mail: gaby@scilearn.com
[b]Corresponding author's address: AudioLogic, 4870 Sterling Dr., Boulder, CO 80301. Electronic mail: jim@audiologic.com

anechoic conditions to 3.1 dB in a simulated conference-room environment. The subject improvements in SRT ranged from 26.1 to 3.6 dB under the same conditions. Thus the improvement in weighted SNR is directly related to a comparable improvement in speech intelligibility in noise for a variety of array designs and signal-processing techniques.

Two additional studies have evaluated the advantage of directional microphones for different levels of reverberation and/or processing strategies. Kompis and Dillier (1994) tested both normal-hearing and hearing-impaired listeners with isolated words. They used a microphone array consisting of a single cardioid microphone over each ear, and combined the two signals using an adaptive algorithm. A single source of interference was introduced at an angle of 45°. They showed improvements in intelligibility scores of approximately 18%. Hawkins and Yacullo (1984) compared the SNR for 50% word recognition for a single omnidirectional microphone with that for a single directional microphone. They found improvements of about 3 to 4 dB for the directional microphone in the presence of a single source of noise introduced behind the listener's head in a reverberant room.

The strongest evidence for the benefit of microphone arrays, however, would require acoustic tests of the arrays in representative rooms and measurements of speech intelligibility for hearing-impaired subjects in the same rooms. A set of acoustic tests was performed by Kates and Weiss (1996). They used a linear array consisting of five omnidirectional microphones having a uniform 2.5-cm spacing to acquire data in an office and in a conference room. The data was then processed using delay-and-sum beamforming, oversteered delay-and-sum (Cox et al., 1986), superdirective (Cox et al., 1986; Kates, 1993), and adaptive array processing (Cox et al., 1987). All of the processing techniques were applied off-line to the same set of microphone signals. The average results for a single source of interference and for a combination of five noise sources were within 1 dB of each other, as were the average results for the adaptive and superdirective arrays. The superdirective array gave an AI-weighted array gain of 9 dB as compared with 5 dB for the delay-and-sum beamforming. This latter figure is very close to the theoretical result of 4.8 dB for a cardioid microphone in free space.

The primary objective of the present study is to determine the degree to which the array used by Kates and Weiss (1996) improves speech intelligibility. A secondary objective is to determine the extent to which the acoustic measurements made by Kates and Weiss correlate with those obtained from subject testing. If the two sets of data are highly correlated, then the time and costs involved with human subject testing can be avoided. Two of the four approaches used by Kates and Weiss (1996) were evaluated here: delay-and-sum beamforming and superdirective processing. These two processing techniques were chosen because delay-and-sum beamforming and superdirective processing represent the simplest analog and digital systems, respectively, and because the acoustic measurements for the superdirective and adaptive processing systems were so similar that subject testing would be unlikely to differ. A single omnidirectional microphone and a single cardioid microphone were selected

as reference conditions to provide a comparison with current hearing-aid technology. Two rooms, an office and a conference room, were used for the data acquisition in order to determine the effects of room size on processed speech intelligibility.

Speech intelligibility in noise was assessed with two different measurement techniques. A speech reception threshold (SRT) was measured using 11 spondees presented in isolation (Dubno et al., 1984), and a speech intelligibility rating (SIR) was measured using continuous discourse (Cox and McDaniel, 1984, 1989; Speaks et al., 1994). The SRT measurement was chosen to permit comparison with other measurements of microphone array performance in rooms (Hoffman et al., 1994). However, since isolated words minimize the effects of room reverberation on speech intelligibility, sentence material was also used. The sentence material allowed a direct comparison with the results of Soede et al. (1993b), and the use of these two different intelligibility measures in the array evaluation also permitted a comparison of objective and subjective assessment of the same set of microphone-array processing techniques in a reverberant room. While these speech materials emphasize the low-frequency portions of the speech spectrum (Van Tasell and Yanz, 1987), they are representative of actual listening conditions under which hearing-impaired listeners would evaluate processing benefits. Stimuli that emphasize the high frequencies of speech, such as nonsense syllables, could lead to different results.

I. METHOD

A. Data acquisition

The array used for experiments was 10 cm long and consisted of five uniformly spaced Knowles EK-3033 omnidirectional microphones. The array was mounted in an endfire orientation. This is the same array as used by Kates and Weiss (1996) for their acoustic measurements. The single omnidirectional reference microphone was the rearmost omnidirectional microphone in the array. The single cardioid microphone was a Knowles EL-3083 mounted 2.5 cm behind the omnidirectional reference microphone. The microphone outputs were bandlimited to 5 kHz and were sampled at 10 kHz.

Two rooms, an office and a conference room, were used for the measurements. The array and reference microphones were placed at a height of 1.4 m above the floor using a floor stand placed near the center of each room. The office had a reverberation time of 250 ms, and the conference room had a reverberation time of 600 ms. The critical distance (distance from the sound source for which the steady-state direct and reverberant sound powers are equal) in both rooms was about 1 m, and the direct-to-reverberant ratio for the speech stimuli was about -6 dB in the office and -10.5 dB in the conference room. Dimensions and layouts for the two rooms are given by Kates and Weiss (1996) along with related acoustic parameters.

Stimuli were presented one at a time over a loudspeaker with the microphone responses sampled and stored on the computer for later processing. Speech stimuli were presented

at an azimuth of 0°, and the noise stimuli were presented at azimuths of 60°, 105°, 180°, 255°, and 300°, counterclockwise around the array. The uncorrelated noise stimuli at each loudspeaker position consisted of multi-talker speech babble. The combined noise source used for the intelligibility tests was formed by summing at equal intensities the babble signals from the five noise azimuths; this combination produced a diffuse noise field of the sort that would be found in a restaurant or similar environment where several people are talking simultaneously.

Signal level calibration was carried out by playing the speech babble through each loudspeaker in turn, and adjusting the amplifier gain to give a sound pressure of 85 dB SPL at the microphone array position. Because the calibration level included both the direct and reverberated sound, the amount of direct sound in the speech stimuli at the array position differed in the two test rooms. Furthermore, the recordings of the SRT and SIR source materials were made at different levels, with the result that the two sets of stimuli were reproduced at different intensities in each room since the same amplifier gain settings were used for both sets of stimuli. These differences in stimulus and room characteristics mean that absolute SNR values can not be accurately compared across stimuli, and may differ in the two rooms for comparable speech intelligibility. Relative comparisons normalized by the reference omnidirectional microphone measurements, however, are valid across both stimuli and rooms.

B. Array processing

The array processing was implemented using the block frequency-domain approach described by Kates and Weiss (1996). The equivalent of 16-tap filters at each microphone were used. The filter weights for the delay-and-sum beamforming are given by $\mathbf{w}(k) = \mathbf{d}(k)/M$, where $\mathbf{d}(k)$ is the steering vector (vector giving the phase shift from one microphone to the next as a wave arriving from 0°, propagates across the array) for fast Fourier transform (FFT) frequency index k, and M is the number of microphones in the array. Delay-and-sum beamforming optimizes the array performance under the assumption that the dominant source of noise is the self-noise of the microphones and not the ambient noise field.

The weights for the superdirective processing optimize the array performance for an assumed preexisting ambient noise field. The set of microphone weights in each FFT frequency bin is chosen to optimize the array output SNR subject to a constraint that a signal from the endfire direction be passed with unit gain. The superdirective weights at each frequency are given by (Monzingo and Miller, 1980)

$$\mathbf{w}(k) = \frac{\mathbf{R}^{-1}(k)\mathbf{d}(k)}{\mathbf{d}^*(k)\mathbf{R}^{-1}(k)\mathbf{d}(k)}, \tag{1}$$

where the asterisk denotes conjugate transpose. The superdirective processing is based on the correlation matrix $\mathbf{R}(k)$ calculated at each frequency for an ideal spherically isotropic noise field (Cheng, 1971). The magnitude of the weight vector was constrained to limit the sensitivity to sensor mis-

alignment (Cox *et al.*, 1986, 1987) using the iterative procedure of Kates and Weiss (1996).

The signal from the reference omnidirectional microphone was used directly. The signal from the reference cardioid microphone was equalized in frequency response and level to more closely match the signal from the omnidirectional microphone. The frequency-response equalization consisted of a boost of 12 dB below 700 Hz, decreasing by −6 dB/oct to 0 dB at and above 2800 Hz. An overall boost of 5 dB was also provided to the output of the cardioid microphone.

Directional patterns for the cardioid microphone and the two array-processing techniques are presented in Fig. 1. The directional pattern for the cardioid microphone in Fig. 1(a) is independent of frequency up to about 4 kHz. There is a null in the pattern at 180°, and the microphone primarily suppresses sound coming from the rear. The delay-and-sum directional patterns in Fig. 1(b) for the array of omnidirectional microphones show that this system has very little array gain at low frequencies. As the frequency increases the main lobe of the array becomes narrower and the sidelobe level decreases, and both effects lead to increased rejection of off-axis noise. The same general pattern is visible in Fig. 1(c) for the superdirective processing. However, the superdirective array offers much better directional behavior at all frequencies than is possible with the delay-and-sum beamforming.

C. Acoustic performance metrics

Two acoustic performance metrics have been proposed for arrays being used for hearing aids. One metric is the Articulation-Index (AI)-weighted array gain (Stadler and Rabinowitz, 1993; Kates and Weiss, 1996). The AI-weighted array gain is calculated from the array gain computed at each frequency of the Fourier-transformed data, and the array gains are combined using weights for each frequency band derived from the Articulation Index importance function given by Kryter (1962). The array gain (Cox *et al.*, 1987) for the kth FFT bin is given by

$$G_A(k) = \frac{|\mathbf{w}^*(k)\mathbf{d}(k)|^2}{\mathbf{w}^*(k)\mathbf{Q}(k)\mathbf{w}(k)}, \tag{2}$$

where $\mathbf{Q}(k)$ is the noise-alone correlation matrix normalized so that $\text{Tr}[\mathbf{Q}(k)] = M$, the number of microphones in the array. The array gain depends on the array weights and on the spatial distribution of the noise, but is independent of the actual signal and noise powers. An array consisting of a single omnidirectional microphone has an array gain of 1. The estimated noise-alone correlation matrix used in the array-gain calculation was smoothed using a low-pass filter having a time constant of 500 ms. Since both the speech and noise were measured in reverberant rooms, this metric gives the ratio of the power in the direct portion of the speech signal to the total direct-plus-reverberant noise power at the array output, normalized by the SNR at the array input. This measure thus represents the directional gain of the array in the noise field.

The AI-weighted array gain is then given by

FIG. 1. (a) Directional pattern for a cardioid microphone. (b) Directional patterns for the five-microphone array using delay-and-sum beamformings. (c) Directional patterns for the five-microphone array using superdirective processing.

$$G_{AI} = \sum_{k=0}^{K} a(k)[10 \log_{10} G_A(k)] \text{ dB}, \qquad (3)$$

where the set of weights $\{a(k)\}$ is the AI importance function weights given by Kryter (1962) reinterpolated for the FFT band edges. Spread of masking effects are ignored in this metric. The AI-weighted array gain G_{AI} is expressed in dB *re*: the array gain for a single omnidirectional microphone.

The array gain given in Eq. (2) differs from the ratio of array output SNR to input SNR that has also been proposed as the basis of the performance metric (Greenberg *et al.*, 1993; Hoffman *et al.*, 1994). The array output SNR is the ratio of the total speech power to the total noise power at the output of the array. The speech and noise powers both include the reverberated as well as the direct components. The ratio of array output to input SNR is given by

$$\Delta \text{SNR} (k) = \frac{\mathbf{w}^*(k)\mathbf{P}(k)\mathbf{w}(k)}{\mathbf{w}^*(k)\mathbf{Q}(k)\mathbf{w}(k)}, \qquad (4)$$

where $\mathbf{P}(k)$ is the speech-alone correlation matrix normalized to have $\text{Tr}[P(k)] = M$ and $\mathbf{Q}(k)$ is the normalized noise-alone correlation matrix. The processing benefit using this metric is then summed over frequency using the AI weights. Greenberg *et al.* (1993) defined the weighted summation over frequency as the intelligibility-weighted change in SNR, which they denoted by G_I.

The two metrics differ in how the room reverberation is treated. In an anechoic environment they yield the same result. In reverberation, however, the metric based on the SNR decreases as the amount of reverberation increases because the array directivity reduces the amount of off-axis reverberated speech passed by the array. Thus the two metrics can differ substantially under conditions of negative (in dB) speech direct/reverberant ratio as was used in the present experiment. Kates and Weiss (1996) present more discussion on this topic.

D. Subjects

Eighteen males aged between 67–77 years took part in the study (mean=73.1 yr, standard deviation=3.1 yr). All had symmetrical sensorineural hearing losses, and had worn hearing aids between 3 and 28 years. Two of these 18 subjects, however, were unable to carry out the SIR test satisfactorily, and so their data was excluded from the study. Figure 2 shows the audiometric configurations of the remaining 16 individuals whose data were analyzed.

E. Procedures
1. General procedures

Subjects came for testing on two occasions. At visit one, pure tone audiometry, most comfortable level (MCL) determination, and speech reception threshold (SRT) measurement took place. At visit two the speech intelligibility ratings (SIR) were made. Details of the SRT and SIR methodology
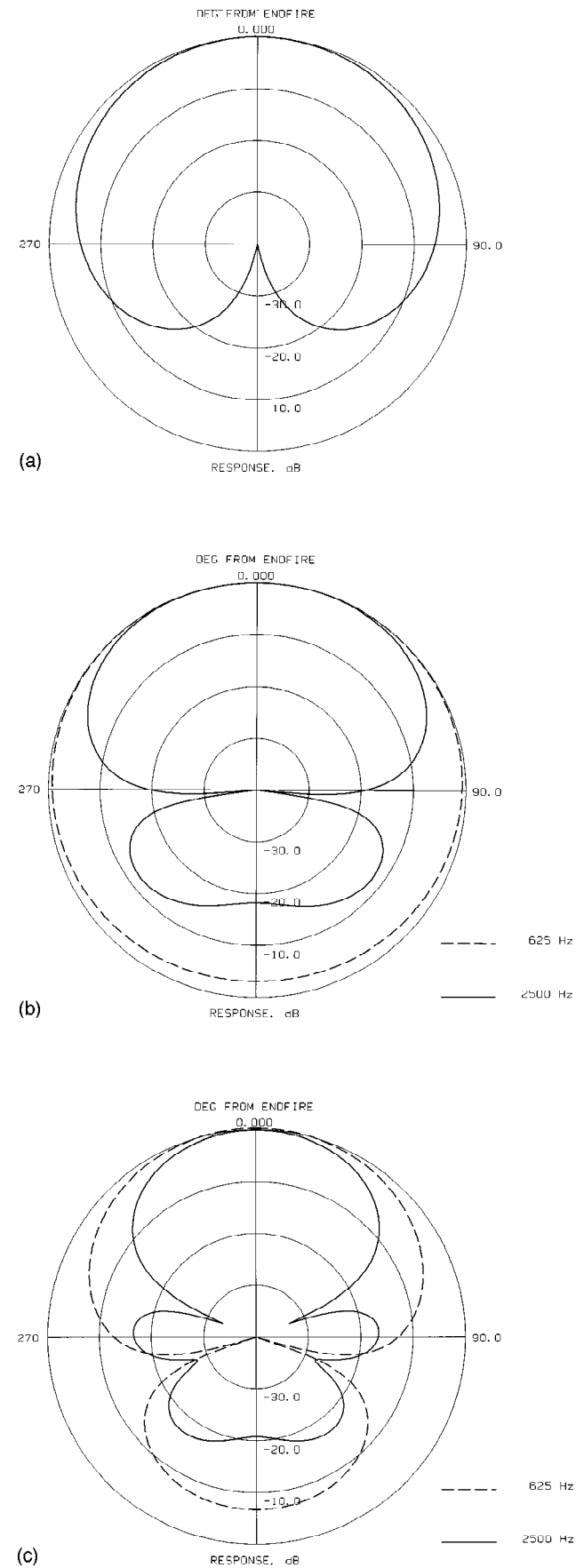
G. H. Saunders and J. M. Kates: Hearing-aid array processing
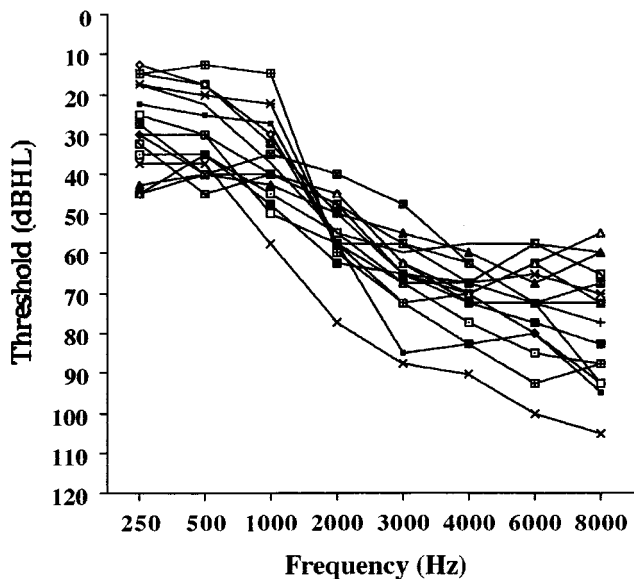
FIG. 2. Audiograms for the 16 hearing-impaired subjects used in the experiment.

are described below. To prevent learning effects influencing the group data, the order in which subjects carried out the eight test conditions (4 processing conditions×2 rooms) was determined by an 8×8 Latin square design. Two subjects followed each of the eight randomized condition orderings.

All testing was carried out in an Acoustic Systems double-walled test booth. Stimuli were presented diotically through TDH-50P headphones. The speech signal was played at a level midway between upper and lower MCL. To determine MCL, a variation of the method of Hawkins *et al.* (1987) was employed. Subjects listened to an unprocessed passage of speech that was equated in output level to the processed speech stimuli. Subjects were asked to rate its loudness on a three-point scale: ''comfortable but slightly loud,'' ''comfortable,'' or ''comfortable but slightly soft.'' The level of the speech was altered in 2-dB steps in an up–down method until two responses of ''comfortable'' were given at the same level. Midpoint MCLs ranged from 78–94 dB SPL, with a mean of 85 dB SPL. The signal presentation at MCL was chosen to be representative of real-world hearing-aid use where the listener has access to a volume control. The use of MCL, as opposed to a gain fixed by the experimenter, could reduce the audibility of some high-frequency speech sounds if the subject preferred a setting having reduced gain.

All speech and noise stimuli were stored in their processed form on the hard drive of an IBM-compatible 386 computer. The levels of the speech and noise were controlled independently by a ProAudio Spectrum-16 computer board. The signals were then mixed and passed through a TEAC EQA-220 ten-channel graphic equalizer that had been set to shape the output to the National Acoustics Laboratories (NAL) prescribed frequency response for each subject (Byrne and Dillon, 1986). The shaped signal was amplified by a Rane SM26 splitter mixer and a GSI 10 audiometer before being played at the headphones. All calibration and the setting up of the NAL response was done with a sound level meter at the headphones.

2. SRT

The SRT for 50% correct performance was measured using a subset of 11 W-1 spondees, known to be highly homogeneous in intelligibility (Dubno *et al.*, 1984). The same set of spondees was used to assess adaptive array performance by Hoffman *et al.* (1994). The spondees were: *drawbridge, eardrum, grandson, northwest, padlock, playground, railroad, sidewalk, toothbrush, woodwork*, and *workshop*. They were played at MCL in the presence of background noise that varied adaptively in level. The subjects' task was to repeat back each word after hearing it. They were given a written list of the 11 spondees and encouraged to use the list when deciding upon their responses. An up–down adaptive procedure was used to estimate 50% correct performance (Levitt, 1971). Testing began at a +20 dB SNR; the noise level was increased when a spondee was repeated correctly and decreased if repeated incorrectly. The initial step size was 4 dB, which was reduced to 1 dB following the second reversal. Testing was continued until 16 reversals had occurred. The SNR at the final ten reversals was averaged for the SRT. Testing continued for a minimum of three runs per condition, or until two SRTs were obtained that were within 1 dB of each other. Of the total 128 times the SRT was run (4 processing conditions×2 rooms×16 subjects), this criterion was met in 3 runs on 66 occasions, in 4 runs on 55 occasions, and in 5 runs on 7 occasions. The final SRT for each condition was calculated by averaging the best two SRTs from above. Each individual SRT took about 2 min to obtain.

3. SIR

Speech intelligibility ratings (SIR) were made using the sentence materials described by Beck and Speaks (1993). Subjects heard passages of everyday speech spoken by a male in the presence of background noise. The passages were originally developed by Cox and McDaniel (1984, 1989), and were further equated for intelligibility by Speaks *et al.* (1994). Even though spondees and continuous discourse yield similar speech intelligibility results under nonreverberant laboratory conditions (Dubno *et al.*, 1984), there could still be differences in intelligibility in a room due to the reverberant tail of one sound interfering with the intelligibility of a subsequent sound in the sentence. The continuous discourse thus provided a test for additional speech understanding difficulties caused by the reverberation.

Twenty-four SIR passages were used for testing, with a further six used for practice. The purpose of the test was to determine the SNR at which the subjects judged they could understand 50% of the spoken passage. For the test, the subjects were seated in the test booth, and in front of them was a joystick and monochrome monitor. In order to begin a run they pressed a button on the joystick. The speech and noise were then played at a highly favorable (+20 dB) SNR. The subjects' task was to alter the level of the noise by moving the joystick upwards (to increase the noise level) or down-

wards (to decrease the noise level) until they judged that they could "just understand half of what was being said." It was emphasized that they should neither understand more nor less than half. When subjects determined they had reached the appropriate SNR they pressed the joystick button once more. The computer then stored the final SNR in an output file, and restarted the test procedure. Subjects were provided with a visual set of instructions on the monochrome monitor and feedback in the form of an illuminated arrow that indicated whether they were raising or lowering the level of the noise. The test passages were approximately 45 s long. Subjects were encouraged to utilize all of this time, by adjusting the noise level above and below their threshold, in order to obtain the best judgment of 50% understanding that they could. On average subjects reached threshold after about 30 s. Testing in each condition continued until two SIRs were obtained that were identical, with an additional SIR within 1 dB of these. The modal number of runs required to meet this criterion was four, with a range of three to eight. The final SIRs were calculated by averaging the SNRs for the three SIR runs used to meet the criterion. Each SIR took approximately 30 s to obtain.

Prior to testing, subjects were trained in the test procedure with the six unprocessed passages. If after six runs performance was not consistent, training continued until it was. No feedback was provided since the objective of the test was to obtain a subjective rating of speech intelligibility. As mentioned above, 2 of the original 18 subjects were unable to give intelligibility judgments that converged to a stable SNR estimate; all other subjects had no difficulty obtaining consistent results.

II. RESULTS

The absolute SNRs computed for the two speech intelligibility measures differ as a result of the signal level calibration. Therefore, repeated-measures analyses of variance were carried out on the data from each speech measure separately, using room and processing conditions as within-subject factors. For both test measures there were significant main effects of room [SRT: $F(1,15) = 198.1$, $p < 0.001$; SIR: $F(1,14) = 205.0$, $p < 0.001$] and processing condition [SRT: $F(3,45) = 70.6$, $p < 0.001$; SIR: $F(3,42) = 96.8$, $p < 0.001$] and there was a significant interaction of room×processing condition for the SIR [$F(3,42) = 5.4$, $p < 0.01$]. Tukey HSD *post hoc* tests ($p < 0.01$) show that performance on both measures was significantly better in the office than in the reverberant conference room and that performance in each processing condition differed significantly from the others ($p < 0.05$). Figures 3 and 4 show the mean scores and 95% confidence intervals for performance on both tests, broken down by room and processing condition. Figure 5 shows the results of *post hoc* analyses for the room×condition interactions. For both measures it is seen that subjects performed significantly better with the superdirectional processing in the office than in all other conditions. Furthermore, superdirectional processing improved performance in the reverberant conference room to the extent that it became better than performance with the omnidirectional processing in the less-reverberant office environment.
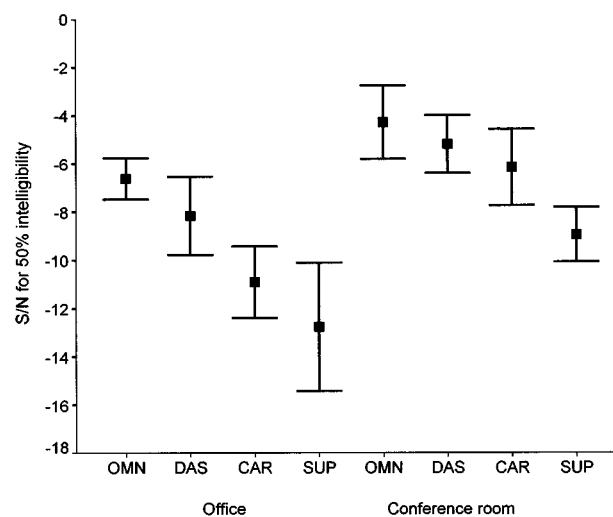
FIG. 3. Signal-to-noise ratio (SNR) required for 50% speech intelligibility and 95% confidence intervals for the SRT spondees. Data are presented for the omnidirectional microphone (OMN), delay-and-sum beamforming (DAS), cardioid microphone (CAR), and superdirective array processing (SUP) in both the office and the conference room.

In order to determine whether the SRT and SIR test measures provided different information, a comparison was made by looking at the performance for the cardioid microphone and arrays within each room relative to the performance for the omnidirectional microphone in the same room. These normalized data are presented in Table I, where the SNR for the omnidirectional microphone has been subtracted from the SNR values for the other systems within the given room. The processing benefit is the increase in background noise level that can be tolerated while still achieving 50% intelligibility. The normalized scores for the SRT and SIR tests are quite similar for both rooms. A repeated-measures ANOVA was carried out to further investigate this similarity. The main effect of test measure was not significant, nor were the two-way interactions of test measure×room and test
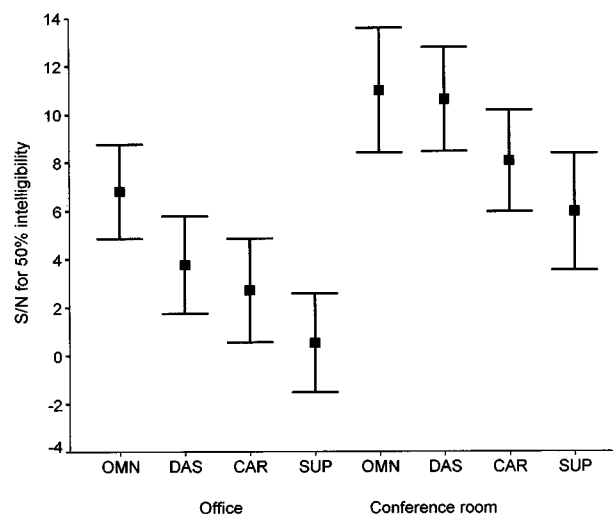
FIG. 4. Signal-to-noise ratio (SNR) required for 50% speech intelligibility and 95% confidence intervals for the SIR passages. Data are presented for the omnidirectional microphone (OMN), delay-and-sum beamforming (DAS), cardioid microphone (CAR), and superdirective array processing (SUP) in both the office and the conference room.
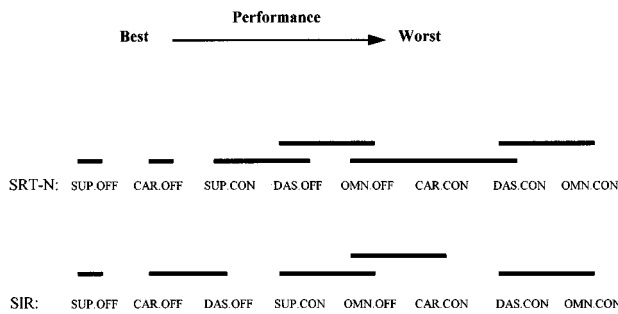
FIG. 5. Processing and room conditions giving similar performance. The bars indicate results that were not significantly different from each other at the $p = 0.05$ level using the Tukey HSD *post hoc* analysis. The processing is the omnidirectional microphone (OMN), delay-and-sum beamforming (DAS), cardioid microphone (CAR), and superdirective array processing (SUP). The rooms are the office (.OFF) and the conference room (.CON).

measure×processing condition. The three-way interaction of test measure×room×processing, however, was significant $[F(3,42) = 2.9, p < 0.05]$. *Post hoc* tests ($p < 0.05$) show that subjectively individuals judged the delay-and-sum processing to be more beneficial in the office than their SRT performance suggested.

III. DISCUSSION

A. Frequency effects

The comparison of processing techniques in Table I shows a very small benefit for the array of omnidirectional microphones using delay-and-sum processing, and a larger benefit for the cardioid microphone. As shown in Fig. 1(b), the delay-and-sum processing has very little directional selectivity at low frequencies, where most of the speech power is concentrated (Dunn and White, 1940), and develops much better directional characteristics at high frequencies. On the other hand, Fig. 1(a) shows that the cardioid microphone has a much better directional pattern at low frequencies than that of the array using delay-and-sum beamforming. It thus appears that the low-frequency array performance dominated the speech intelligibility results.

The superdirective array benefit for speech intelligibility is noticeably lower than the 9 dB found using the AI-weighted array gain performance metric for the same array and rooms (Kates and Weiss, 1996). However, the relative difference in performance between the superdirective and delay-and-sum array processing is consistent with the acous-

TABLE I. Benefit for the different processing approaches in the two test rooms relative to the performance for the omnidirectional microphone in the same room.

	SRT (dB)	SIR (dB)
Office		
Delay-and-sum	1.5	3.0
Cardioid	4.3	4.0
Superdirective	6.2	6.3
Conference room		
Delay-and-sum	0.9	0.5
Cardioid	1.9	3.0
Superdirective	4.7	5.1

tic measurements; the superdirective array processing gave on average 5.6 dB better performance than the omnidirectional microphone, and 4.1 dB better performance than the delay-and-sum beamforming. The superdirective array processing also results in much better directivity at high frequencies than at low frequencies, as shown in Fig. 1(c).

The difference between the acoustic measurements and the intelligibility results for the superdirective array processing could be due to a greater importance of the low-frequency array gain for the subjects and test materials than is represented by the AI weights (Kryter, 1962) derived for normal-hearing subjects listening to isolated syllables. The AI weights for hearing-impaired individuals tend to give less importance to the high frequencies of speech than observed in the weights for normal-hearing individuals (Pavlovic *et al.*, 1986). The AI weights for continuous discourse also give less importance to the high frequencies of speech than the weights determined for nonsense syllables (Studebaker *et al.*, 1987; Pavlovic, 1987). Thus the combination of hearing-impaired subjects listening to continuous discourse could greatly increase the importance of the low frequencies of speech relative to that typically assumed for normal-hearing subjects.

A further explanation for the difference between the acoustic measurements and the intelligibility results could be the masking caused by the interfering noise (Egan and Hake, 1950), especially since hearing-impaired listeners are more susceptible to the effects of masking than are normal-hearing subjects (Glasberg and Moore, 1986; Gagné, 1988). The directional characteristics of the arrays vary with frequency, and therefore modify the spectrum of the interfering speech babble. The array gain increases with increasing frequency, providing a greater reduction in the amount of off-axis noise power passed by the array and resulting in a low-pass filtering of the interference spectrum after the array processing. The residual noise at low frequencies could then mask high-frequency speech components.

B. Reverberation

A second concern is the effect of reverberation on speech intelligibility in rooms. The results presented in Table I indicate that the subject performance in the conference room was 1.5 dB worse than the performance in the office when normalized by the performance for the omnidirectional microphone in each room. This room difference occurs for the cardioid microphone as well as for the arrays, and is apparent for both the SRT and SIR tests. The direct/reverberant ratio for the speech loudspeaker in the conference room was estimated to be -10.5 dB, as opposed to -6 dB in the office. Since the level of the direct-plus-reverberated speech signal at the array position is the same for the two rooms, the direct sound is less intense and the speech reverberation more intense in the conference room than in the office.

The difference in the direct-to-reverberant speech ratio may account for the difference in array benefit in the two rooms. The array directional pattern maintains a constant gain in the direction of the speech source, but decreases the gain for other directions. Thus the intensity of the reverber-

ated speech is reduced by the array along with the noise. If the reverberation contributes to speech intelligibility, then the array directionality will reduce intelligibility by the degree to which the useful reverberated speech power is attenuated. The assumption made in the weighted SNR performance metric is that all of the speech reverberation is beneficial. The AI-weighted array gain metric, on the other hand, ignores the speech reverberation and predicts that there will be essentially no difference in performance between the two rooms given the same direct speech signal power.

The data in Table I are consistent with the early reflections in the room providing some benefit for intelligibility, but not the later reflections. Lochner and Burger (1964), in experiments using normal-hearing subjects, found that binaural speech intelligibility in the presence of a single reflection was greater than that for the direct signal alone for reflection delays up to 30 ms. The intelligibility decayed to that of the direct signal alone as the delay was increased beyond 30 ms. Nábělek and Robinette (1978) found similar results for both normal-hearing and hearing-impaired subjects, with a reflection having a delay at or below 20 ms contributing to intelligibility and with longer delays providing reduced benefit or some possible reduction in intelligibility. The benefits for the early reflections in speech intelligibility were consistent with combining the power in the reflection with that in the direct sound. Thus including early reflections, but not late reflections, along with the direct speech sound in the speech power used in the performance metrics may lead to more accurate predictions of array benefit for speech intelligibility for both normal-hearing and hearing-impaired listeners.

C. Performance metric

A new room performance metric was computed, taking into account both the benefit of early room reflections and the deleterious effects of masking. The contribution of the early reflections to the useable speech power was estimated by forming a weighted average of the array gain and ΔSNR performance measures. Both metrics include the direct speech signal, but only the ΔSNR includes the speech reverberation. The contribution of the early reflections can not be separated from that of the late reflections in the ΔSNR, since it is a steady-state measurement, but the amount of power in the early reflections can be approximated as a fraction of the total speech output power. The weighted average, when fit to the subject data from the two rooms given in Table I, results in the room gain given by $G_R(k) = 0.67 G_A(k) + 0.33 \Delta \text{SNR}(k)$. Combining Eq. (2) with Eq. (4) results in the expression for the room gain:

$$G_R(k) = \frac{0.67 |\mathbf{w}^*(k)\mathbf{d}(k)|^2 + 0.33 \mathbf{w}^*(k)\mathbf{P}(k)\mathbf{w}(k)}{\mathbf{w}^*(k)\mathbf{Q}(k)\mathbf{w}(k)}. \quad (5)$$

The second step in computing the array benefit is to add the effects of upward spread of masking. The Articulation Index calculation (Kryter, 1962) contains a procedure for estimating the amount of masking of speech by noise for normal listeners. Greenberg *et al.* (1993) include masking in

TABLE II. Comparison of array acoustic performance metrics. The room metric is a weighted average of the array gain and ΔSNR at each FFT frequency bin, to which a masking model and AI weights are then applied.

	Average of SRT and SIR (dB)	AI-weighted array gain (dB)	AI-weighted ΔSNR (dB)	Room metric (dB)
Office				
Delay-and-sum	2.3	5.0	3.0	2.4
Superdirective	6.3	9.5	6.5	6.2
Conference room				
Delay-and-sum	0.7	5.0	1.2	1.2
Superdirective	4.9	9.0	3.0	4.7

their general definition of the intelligibility-weighted change in SNR, but masking has not been included in the actual AI-weighted array gain (Stadler and Rabinowitz, 1993; Kates and Weiss, 1996) or intelligibility-weighted SNR (Greenberg and Zurek, 1992; Hoffman *et al.*, 1994) metrics used to evaluate microphone array performance. The AI-weighted room gain calculation was therefore modified to include upward spread of masking for the group of hearing-impaired listeners. The masked spectrum model consists of estimating the incoming noise power in each frequency band for bands contributing equally to speech intelligibility, extending a line at -3 dB *re*: the band level upwards in frequency for 200 Hz or 1/2 octave, whichever is greater, and then continuing upwards in frequency with a line having a slope of -8 dB/oct. If the masked spectrum in dB is greater than the measured noise spectrum in any frequency band, the room array gain in dB in that band is reduced by the their difference. The parameters of the masked spectrum model were chosen to match the subject data presented in Table I.

The results of using the performance metric incorporating the room gain and the upward spread of masking are presented in Table II, along with an average of the improvements obtained on the SRT and SIR from Table I and the AI-weighted ΔSNR and array gains without masking. It is seen that the AI-weighted array gain overestimates the overall level of performance, but gives accurate performance differences between the delay-and-sum and superdirective processing. The weighted ΔSNR, on the other hand, reproduces the office intelligibility test results very closely, but underestimates the overall superdirective processing array benefit in the conference room as well as the difference between the delay-and-sum and superdirective processing results. When the early reflection power masking model is incorporated into the AI-weighted room gain, it gives results very close to the intelligibility test results in the office and similar results to those for the weighted ΔSNR metric. In the conference room, the room gain and masking model gives a much more accurate fit to the intelligibility test results than either the AI-weighted array gain without the masking model or the weighted ΔSNR. While the amount of data being fitted is quite limited, the success of the masking model in fitting the observed data for both array-processing techniques in both rooms suggests that upward spread of masking should be considered in designing microphone arrays and in measuring their performance.

D. Subjective versus objective tests

Test results obtained with the SRT did not differ significantly from those obtained on the SIR. Both the SRT and SIR stimuli emphasize the low frequencies of the speech spectrum. However, the spondees in the SRT were presented to subjects without a carrier phrase, which would have minimized the interfering effects of reverberated speech, while the connected discourse of the SIR would have maximized the interference of the reverberated speech upon subsequent words in the passage. The finding that test results were equivalent for both sets of materials indicates that the lingering reverberation of previous words was not an important factor in speech identification for the rooms used in this study, probably because by the time the reverberant tail for a given word was overlaid on the following word, the intensity of the reverberation had become relatively low. The ongoing reverberation of a word, however, can still interfere with speech sounds occurring later in the same word; this effect of reverberation would be essentially the same for the isolated spondees as for the connected discourse.

A single run on the SIR test took approximately 30 s, while a single run of the SRT took approximately 2 mins. However, once the time taken for instruction and training on the two tests is included, testing time for the measures was about the same. Since both measures yield essentially the same information, both would be candidates for use as a clinical tool for adjusting hearing-aid parameters under realistic reverberation conditions. However, the SRT would probably be the more suitable measure because self-adjusted tests like the SIR can be affected by nonauditory factors (Saunders *et al.*, 1992) and because some subjects were unable to estimate intelligibility reliably using the SIR technique.

E. Related experiments

The results of this experiment compare favorably with the array results reported for other subject experiments once the differences in test conditions and array design and processing are accounted for. The average benefit of 3.3 dB relative to the omnidirectional microphone found for the cardioid microphone is consistent with the 3–4 dB measured by Hawkins and Yacullo (1984) for a sound source in front and the noise source behind the head in a reverberant room. Soede *et al.* (1993a) made acoustic measurements for an array of five uniformly-spaced cardioid microphones with delay-and-sum beamforming and measured improvements in the range of 5–12 dB. Using this same array, subject SRT performance for sentence materials improved by an average of 7 dB (Soede *et al.*, 1993b). In this study subjects' performance improved by 1.5 dB for delay-and-sum beamforming with an array of five omnidirectional microphones. However, the beam pattern for an array of directional microphones is the product of the beam pattern for the array of omnidirectional microphones with that for the directional microphones. The multiplication of the beam patterns corresponds roughly to adding the SNR or SRT benefits in a diffuse field; the 1.5 dB for the omnidirectional delay-and-sum array plus the 3.3 dB for the cardioid microphone yields 4.8 dB, which is close

to the benefit found by Soede *et al.* (1993b). The remaining performance differences may well be due to the differences in the stimuli and interfering noise fields.

The performance of an adaptive system will converge to that of a superdirective system in a diffuse noise field. A small number of noise sources (fewer than $M-1$ sources for M microphones in the array) creates a situation that favors the adaptive system. Using a single source of interference, Hoffman *et al.* (1994) found improvements in SRT of 10.4 dB in a simulated living room and 3.6 dB in a simulated conference room for a seven-microphone head-mounted adaptive array, and improvements of 7.3 dB in the living room and 2.2 dB in the conference room for a head-mounted three-microphone adaptive array. Averaging the results for the two arrays to approximate the benefit of a five-microphone array gives 8.9 dB for the living room and 2.9 dB for the conference room. The head-mounted array, with microphones positioned around the forehead from ear to ear, is longer than the 10-cm array used in the present study, and also extends in two dimensions instead of the one dimension used for the linear superdirective array.

In comparison with the superdirective array, the adaptive system of Hoffman *et al.* (1994) gave about 2.6 dB better SRT performance in the living room (office), and 2.0 dB worse performance in the conference room. In the simulated living room, the adaptive system is able to steer a spatial response null in the direction of the noise loudspeaker, and therefore gives better results than the superdirective system in the office. In the simulated conference room, however, the reverberation is sufficient to create a much more diffuse noise field, and the performance of the adaptive system is inferior to that of the superdirective system. While these comparisons of fixed-coefficient with adaptive processing are necessarily imprecise due to the differences in test conditions, they suggest that the performance of the superdirective array will be similar to that of an adaptive system in many situations, and may exceed that of an adaptive system in reverberant listening situations. This observation is consistent with the acoustic measurements of Kates and Weiss (1996), who found only a 1-dB advantage for adaptive processing over superdirective processing when applied to the same 10-cm linear array consisting of five uniformly spaced omnidirectional microphones.

IV. SUMMARY AND CONCLUSIONS

This study investigated the use of microphone arrays to improve speech intelligibility in noise for the hearing impaired. The linear array consisted of five omnidirectional microphones and was 10 cm long (about the size of a pen or pencil). The array is thus small enough to be hand-held or built into an eyeglass frame. Delay-and-sum beamforming and superdirective processing strategies were compared with a single omnidirectional and a single cardioid microphone on subject test measures.

The experimental results indicate that the single cardioid microphone is more effective than the array of omnidirectional microphones using delay-and-sum beamforming for the speech materials considered. It was postulated that the relatively poor performance of the delay-and-sum beam-

forming was due to masking of the speech by the residual low-frequency noise passed by the array. The fact that adding a masking component to the AI-weighted room gain acoustic performance metric greatly improved the match between the metric and the subject intelligibility results confirms this interpretation. The greatest benefit was obtained for the array using superdirective processing. It is anticipated that the superdirective performance would be better still if an array of cardioid microphones were used (Soede *et al.*, 1993a; McKinney and DeBrunner, 1993; Kompis and Dillier, 1994), since the array directivity would be increased at all frequencies by the directional elements.

The intelligibility test results showed performance in the nonreverberant office to be better than in the reverberant conference room, reinforcing the observation that increased reverberation leads to decreased speech intelligibility (Duquesnoy and Plomp, 1980; Nábělek, 1982). Modifying the array gain performance metric to give the room array gain, by incorporating an estimate of the power contributed to the direct speech signal by the early reflections, results in a much better match between the subject data and the acoustic performance metric. This demonstrates that accurate predictions of array performance in rooms need to encompass both estimates of the effects of early reflections and effects of frequency-domain masking.

Once the data for each room were normalized to the performance for the omnidirectional microphone, the SRT and SIR tests gave equivalent results for the hearing-impaired subjects. Thus clinicians should be free to use whichever test is more convenient for testing hearing-aid performance in noise and reverberation. However, due to the potential effects of nonauditory factors, the SRT might be a better choice.

The results indicate several areas for continued research. Improved array directional characteristics at low frequencies are needed to reduce the effects of masking on speech intelligibility; the use of cardioid microphones in the array might be one way to achieve this end. Binaural arrays should also be investigated in order to exploit binaural release from masking and thus improve speech intelligibility in noise (Dirks and Wilson, 1969; Markides, 1977). Binaural release from masking complements the array performance since the release from masking is greatest at low frequencies, where the array directivity is poorest. More accurate acoustic performance metrics are also needed, since neither the AI-weighted array gain nor the weighted ΔSNR accurately predicted the performance benefits found in this study. Inclusion of early room reflections and upward spread of masking into the performance metric leads to much more accurate results for the data considered in this paper.

ACKNOWLEDGMENTS

Beck, W. G., and Speaks, C. (**1993**). ''Intelligibility of selected passages from the Speech Intelligibility Rating (SIR) Test,'' J. Speech Hear. Res. **36,** 1075–1082.

Byrne, D., and Dillon, H. (**1986**). ''The National Acoustics Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid,'' Ear Hear. **7,** 257–265.

Cheng, D. K. (**1971**). ''Optimization techniques for antenna arrays,'' Proc. IEEE **59,** 1664–1674.

Cox, R. M., and McDaniel, D. (**1984**). ''Intelligibility ratings of continuous discourse: Application to hearing aid selection,'' J. Acoust. Soc. Am. **76,** 758–766.

Cox, R. M., and McDaniel, D. (**1989**). ''Development of the Speech Intelligibility Rating (SIR) test for hearing aid comparisons,'' J. Speech Hear. Res. **32,** 347–352.

Cox, H., Zeskind, R. M., and Kooij, T. (**1986**). ''Practical supergain,'' IEEE Trans. Acoust. Speech Signal Process. **ASSP-34,** 393–398.

Cox, H., Zeskind, R. M., and Owen, M. M. (**1987**). ''Robust adaptive beamforming,'' IEEE Trans. Acoust. Speech Signal Process. **ASSP-35,** 1365–1376.

Dirks, D. D., and Wilson, R. A. (**1969**). ''The effects of spatially-separated sound sources on speech intelligibility,'' J. Speech Hear. Res. **12,** 5–38.

Dubno, J., Dirks, D., and Morgan, D. (**1984**). ''Effects of age and mild hearing loss on speech recognition in noise,'' J. Acoust. Soc. Am. **76,** 87–96.

Dunn, H. K., and White, S. D. (**1940**). ''Statistical measurements on conversational speech,'' J. Acoust. Soc. Am. **11,** 278–288.

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.

Egan, J. P., and Hake, H. W. (**1950**). ''On the masking pattern of a simple auditory stimulus,'' J. Acoust. Soc. Am. **22,** 622–630.

Gagné, J.-P. (**1988**). ''Excess masking among listeners with a sensorineural hearing loss,'' J. Acoust. Soc. Am. **83,** 2311–2321.

Glasberg, B. R., and Moore, B. C. J. (**1986**). ''Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments,'' J. Acoust. Soc. Am. **79,** 1020–1033.

Greenberg, J. E., and Zurek, P. M. (**1992**). ''Evaluation of an adaptive beamforming method for hearing aids,'' J. Acoust. Soc. Am. **91,** 1662–1676.

Greenberg, J. E., Peterson, P. M., and Zurek, P. M. (**1993**). ''Intelligibility-weighted measures of speech-to-interference ratio and speech system performance,'' J. Acoust. Soc. Am. **94,** 3009–3010.

Hawkins, D. B., and Yacullo, W. S. (**1984**). ''Signal-to-noise ratio advantage of binaural hearing aids and directional microphones under different levels of reverberation,'' J. Speech Hear. Disord. **49,** 278–286.

Hawkins, D. B., Walden, B., Montgomery, A., and Prosek, R. (**1987**). ''Description and validation of an LDL procedure designed to select SSPL 90,'' Ear Hear. **8,** 162–169.

Hoffman, M. W., Trine, T. D., Buckley, K. M., and Van Tasell, D. J. (**1994**). ''Robust adaptive microphone array processing for hearing aids: Realistic speech enhancement,'' J. Acoust. Soc. Am. **96,** 759–770.

Kates, J. M. (**1993**). ''Superdirective arrays for hearing aids,'' J. Acoust. Soc. Am. **94,** 1930–1933.

Kates, J. M., and Weiss, M. R. (**1996**). ''A comparison of hearing-aid array-processing techniques,'' J. Acoust. Soc. Am. **99,** 3138–3148.

Kompis, M., and Dillier, N. (**1994**). ''Noise reduction for hearing aids: Combining directional microphones with an adaptive beamformer,'' J. Acoust. Soc. Am. **96,** 1910–1913.

Kryter, K. D. (**1962**). ''Methods for calculation and use of the Articulation Index,'' J. Acoust. Soc. Am. **34,** 1689–1697.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49,** 467–477.

Lochner, J. P. A., and Burger, J. F. (**1964**). ''The influence of reflections on auditorium acoustics,'' J. Sound Vib. **1,** 426–454.

Markides, A. (**1977**). *Binaural Hearing Aids* (Academic, New York).

McKinney, E. D., and DeBrunner, V. E. (**1993**). ''Directionalizing adaptive multi-microphone arrays for hearing aids using cardioid microphones,'' Proc. 1993 Int. Conf. on Acoust. Speech and Sig. Proc., Minneapolis, MN, 27–30 April 1993, pp. I-177 to I-180.

Monzingo, R. A., and Miller, T. W. (**1980**). *Introduction to Adaptive Arrays* (Wiley, New York), pp. 78–105.

Nábělek, A. K. (**1982**). ''Temporal distortions and noise considerations,'' in *The Vanderbilt Hearing-Aid Report*, edited by G. Studebaker and F. H. Bess (Monographs in Contemporary Audiology, Upper Darby, PA), pp. 51–59.

Nábělek, A. K., and Robinette, L. (**1978**). ''Influence of the precedence effect on words identification by normally hearing and hearing-impaired subjects,'' J. Acoust. Soc. Am. **63**, 187–194.

Pavlovic, C. V. (**1987**). ''Derivation of primary parameters and procedures for use in speech intelligibility predictions,'' J. Acoust. Soc. Am. **82**, 413–422.

Pavlovic, C. V., Studebaker, G. A., and Sherbecoe, R. L. (**1986**). ''An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals,'' J. Acoust. Soc. Am. **80**, 50–57.

Saunders, G. H., Field, D. L. and Haggard, M. P. (**1992**). ''A clinical test battery for obscure auditory dysfunction (OAD): Development, selection and use of tests,'' Br. J. Audiol. **26**, 33–4.2

Soede, W., Berkhout, A. J., and Bilsen, F. A. (**1993a**). ''Development of a directional hearing instrument based on array technology,'' J. Acoust. Soc. Am. **94**, 785–798.

Soede, W., Bilsen, F. A., and Berkhout, A. J. (**1993b**). ''Assessment of a directional microphone array for hearing-impaired listeners,'' J. Acoust. Soc. Am. **94**, 799–808.

Speaks, C., Trine, T. D., Crain, T. R., and Niccum, N. (**1994**). ''A revised speech intelligibility (RSIR) test: Listeners with normal hearing,'' Otolaryngol. Head Neck Surg. **110**, 75–83.

Stadler, R. W., and Rabinowitz, W. M. (**1993**). ''On the potential of fixed arrays for hearing aids,'' J. Acoust. Soc. Am. **94**, 1332–1342.

Studebaker, G. A., Pavlovic, C. V., and Sherbecoe, R. L. (**1987**). ''A frequency importance function for continuous discourse,'' J. Acoust. Soc. Am. **81**, 1130–1138.

Van Tasell, D. J., and Yanz, J. L. (**1987**). ''Speech recognition threshold in noise: Effects of hearing loss, frequency response, and speech materials,'' J. Speech Hear. Res. **30**, 377–386.

Acoustic and physiologic characteristics of inspiratory phonation

Robert F. Orlikoff
Laryngology Laboratory, Head and Neck Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, Box 403, New York, New York 10021

R. J. Baken
Department of Otolaryngology, The New York Eye & Ear Infirmary, New York, New York 10003

Dennis H. Kraus
Head and Neck Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, New York 10021

(Received 21 January 1997; accepted for publication 19 May 1997)

Voice produced on inhaled air is a form of phonation that has received relatively little attention despite its apparent usefulness in the assessment of vocal function. This preliminary investigation was designed to describe the general characteristics of vocalization driven by an ingressive phonatory airflow. Vocal fundamental frequency (F_0), electroglottographic (EGG), and airflow measures were examined in 16 normal men and women, who alternated between inspiratory and expiratory voice. Mean F_0 routinely increased during inspiratory voice segments, shifting on average 5.1 semitones above the subjects' comfortable expiratory voice frequency. EGG data showed inspiratory voice to be associated with a more symmetrical pattern of vocal fold contact characterized by a prolonged interval of increasing contact. Both short-term F_0 variability (jitter) and EGG amplitude perturbation were significantly higher during inspiratory voice. Stroboscopic examination of four of the subjects showed caudal displacement of the larynx and lengthened vocal folds associated with inspiratory phonation. The absolute airflow rate was significantly greater for inspiratory phonation, on average 48.5% higher than during normal expiratory voice. It was also found that both inspiratory pulse and falsetto vibratory patterns could be produced by at least some of the subjects, indicating some control over the mode and frequency of vocal fold vibration when driven by an ingressive airflow. © *1997 Acoustical Society of America.* [S0001-4966(97)02709-4]

PACS numbers: 43.70.Aj, 43.70.Gr [AL]

INTRODUCTION

That voice may be produced on inhaled breath has long been recognized. It is, in fact, not an uncommon form of phonation, usually manifest as a brief audible gasp in vocal expressions of surprise, fear, or joy. It has been documented to occur quite frequently in the cries of infants (Grau *et al.*, 1995). Nevertheless, adult anatomy and physiology is not well-suited to prolonged phonation driven by an inspiratory airflow. The supraglottal airway cannot match the conus elasticus in its ability to funnel a laminar flow toward the glottis. Furthermore, when viewed superiorly, the vocal folds naturally assume a divergent operating point. In theory, this divergence would increase glottal resistance while hindering aerodynamic coupling to the vocal fold mucosa (Gauffin *et al.*, 1983; Scherer and Titze, 1983; Gauffin and Liljencrants, 1988; Scherer and Guo, 1991).

During expiratory voice production, reflexes appear to regulate the duration of vocalization according to the available lung volume (Winkworth *et al.*, 1994, 1995), while phonation itself is associated with the inhibition of the ventilatory response to carbon dioxide (Bunn and Mead, 1971; Mead and Bunn, 1974; Phillipson *et al.*, 1978) and suppression of the cough and gag responses (Davis *et al.*, 1993) that normally work to protect the airway and metabolism. Conversely, reflex activity during the so-called ''prephonatory

inspiratory phase'' (Wyke, 1983) is associated with airway dilatation and the maintenance of patency which would seem to work against glottal sealing. It is not surprising that anecdotal evidence points to phonation on inspired air as generally difficult, uncomfortable, and prone to elicit involuntary coughing.

Because inspiratory (or ''reverse'') phonation has been shown to be associated with distention of the laryngeal ventricles, radiologists have used the maneuver to help distinguish supraglottic and glottic tumors from transglottic disease (Powers *et al.*, 1964; Lehmann, 1965). Otolaryngologists have also reportedly asked their patients to perform this task to aid laryngoscopic examinations in cases where it is difficult to view the vocal folds (Kollár, 1989). Voice therapists have long employed inspiratory voice production in the treatment of ''ventricular (fold) phonation'' and a variety of other voice disorders including puberphonia, spasmodic dysphonia, psychogenic aphonia, and dysphonias due to unilateral vocal fold paralysis, hemilaryngectomy, and benign mass lesions (Kollár, 1973; D'Antonio *et al.*, 1987; Boone and McFarlane, 1988; Shulman, 1993). Still, despite its purported clinical utility in both hyperfunctional and hypofunctional disorders, the nature of inspiratory phonation (IP) has not been studied systematically and is thus only poorly understood.

Timcke *et al.* (1959) analyzed ultrahigh-speed films to investigate the glottal dynamics of an ''easy inspiratory phonation [produced] at low [vocal] intensity.'' They found that, unlike what they had reported as typical for expiratory phonation (EP) (Timcke *et al.*, 1958), glottal width increased relatively slowly and decreased relatively quickly during IP. They also showed that the period of vocal fold approximation was relatively long. Timcke and his colleagues speculated on the basis of the then-current understanding of glottal function that the high speed quotient and low open quotient were due to the fact that, during IP, a relatively high supraglottal pressure works in consort with the Bernoulli force to close the glottis, whereas in EP, a relatively low supraglottal pressure works in opposition to Bernoulli forces. In the years since that investigation, assessment techniques and our understanding of vocal fold vibration and laryngeal aerodynamics have improved significantly, such that a renewed examination of phonation produced using ingressive airflow may yield important insight regarding mucosal and laryngeal airway dynamics. It is the purpose of this preliminary report to describe the general characteristics of inspiratory voice.

I. METHOD

A. Subjects

Eight men, ages 30–44 years (mean 35 years) and 8 women, ages 23–39 years (mean 33 years), served as subjects. None had received voice training and none reported a history of respiratory, vocal or auditory pathology.

B. Experimental procedure

The 2-channel-average output of a model MC2-1 Glottal Enterprises (Syracuse, NY) electroglottograph was used to obtain the subject's electroglottogram (EGG). Electroglottography involves the placement of a pair of surface electrodes on the anterior neck at approximately the level of the vocal folds. Measurement of the transneck impedance to a high-frequency electrical current between the electrodes provides information relating to changes in vocal fold contact area during phonation. The high-pass filtered EGG waveform has also been found to be an effective means of deriving the vibratory, and thus the vocal, period (Horiguchi *et al.*, 1987; Linders *et al.*, 1995; Orlikoff, 1995; Vieira *et al.*, 1997). The EGG electrodes were held in place by a firmly fitting flexible neckband. The linear phase filtering circuitry of the EGG instrument provided 20-Hz high-pass filtering that served to minimize baseline shift and nonvocal EGG influences (Baken, 1992; Vieira *et al.*, 1996).

Phonatory airflow was measured by means of a circumferentially vented pneumotachograph mask and a PTW-1 pressure transducer connected to a Glottal Enterprises MSIF-2 preamplifier which low-pass filtered the flow signal using a 50-Hz cutoff frequency. A pneumotachograph offers exceptional linearity and allows the measurement of both egressive and ingressive airflows (Sullivan *et al.*, 1984; Baken, 1987). However, the use of an anesthesia mask and separate pneumotachograph was rejected in favor of a pneumotachograph mask in order to minimize back pressure and excessive ventilatory dead space. (During pilot testing it was found that the standard mask and coupled pneumotachograph apparatus hindered several potential subjects' ability to produce or to sustain IP.) The airflow system was calibrated separately for each subject using a Glottal Enterprises MCU-2 calibration unit. This unit includes a Gilmont Instruments rotameter connected in series with a specially designed fitted holder for the pneumotachograph-mask assembly. An electret microphone inserted into the central handle of the mask was used to obtain the acoustic signal. Because of acoustic distortion introduced by the mask, the microphone signal was not used for voice analysis, but rather for assistance with the subsequent identification of phonatory segments. The EGG, airflow, and microphone signals were each sampled at 30 kHz using an Alamed (Vega, TX) SIU-2 signal interface and customized Acquire (Ver. 2.0) software running a 12-bit Data Translation DT2821G A/D converter. An IBM-compatible 486DX 66-MHz computer was used for both data acquisition and analysis.

Subjects were seated individually in a sound-attenuated room. Each was instructed to alternate from egressive to ingressive airflow at approximately 2-s intervals while producing the vowel /a/. Reich and his colleagues (1981, 1983) have shown that normal phonatory reaction times do not differ significantly for IP and EP, and the present subjects were told to minimize any break in sound production as they reversed airflow direction. In this way, both expiratory and inspiratory voice ~2-s segments were produced over approximately the same lung volume range. During a minimum of two practice trials the data signals were monitored on a Tektronix model 2212 oscilloscope to assure their quality and to adjust the gain of the amplifiers to make optimal use of the A/D converter's range. Once phonation was begun, simultaneous microphone, EGG, and airflow data were collected for 10 s. Three data trials were obtained at a self-selected comfortable pitch and loudness. Three trials each, at a low-pitched pulse register and at a loft register falsetto pitch, were additionally attempted by each subject after a brief period of training and practice.

Videostroboscopic examinations were conducted on two of the male and two of the female subjects. An EGG-triggered Kay Elemetrics model 9100 Rhino-Laryngeal Stroboscope system and a rigid 70° telescopic oral fiberscope were used to observe the subject's vocal folds during recapitulation of the experimental tasks using the vowel /i/ as per standard endoscopic technique (Watterson and McFarlane, 1991; Faure and Muller, 1992; Hirano and Bless, 1993).

C. Data analysis

Assisted by the acoustic and airflow data, a pair of inspiratory and expiratory segments were extracted from the mid-portion of each subject's voice production. The middle half-second of each of these segments was then subjected to full data analysis. The data reported represent the average across the three productions. Because the subjects used a self-selected comfortable vocal fundamental frequency (F_0), the number of vocal cycles included in the 0.5-s analysis window could be expected to vary accordingly. The average analysis included 59 cycles for the men (range 40–

Electroglottogram
(Vocal Fold Contact Area)

Contact Quotient

$$\frac{A+B}{A+B+C} \times 100$$

Contact Index

$$\frac{A-B}{A+B}$$

Relative Rise Time

$$\frac{A}{A+B+C} \times 100$$
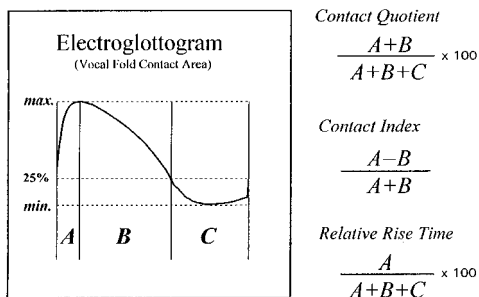
FIG. 1. Measures derived from the subjects' electroglottogram [A = increasing-contact segment; B = decreasing-contact segment; C = minimal vocal fold contact segment; $A + B$ = EGG contact phase].

117) and 123 cycles for the women (range 79–207). All analyses were done using software specifically written for the present experiment.

Each subject's mean vocal fundamental frequency (F_0) and mean jitter were derived from the electroglottogram. Vocal cycles were demarcated by positive going (increasing vocal fold contact) zero crossings of the digitized EGG signal using linear interpolation techniques. Jitter was defined as the average absolute cycle-to-cycle difference in vocal period, divided by the mean period, expressed in percent.

Several measures specific to the EGG vocal fold contact pattern were also obtained (Fig. 1). For purposes of computer synthesis and analysis, Titze (1984, 1989) has defined the abduction quotient as the ratio of the half-distance between the vocal processes to the amplitude of vocal fold vibration. This parameter reflects the degree of medial fold compression along a hypoadducted ''breathy'' to a hyperadducted ''pressed'' (or ''tight'') voice continuum (Scherer and Titze, 1987). In actual practice, however, it is not possible to obtain quick and precise information in vivo about the width of the cartilaginous glottis and the vibratory amplitude. Thus some investigators have turned to the EGG for an indirect assessment of adductory adjustment. The derived measure, the EGG contact duty cycle, usually referred to as the contact quotient (or CQ), is defined as the percent of the entire vibratory cycle wherein vocal fold contact area is greater than some ''minimal level'' (Rothenberg and Mahsme, 1988; Scherer et al., 1988; Orlikoff, 1991; Houben et al., 1992). (In Fig. 1, %CQ = Contact Phase/Vibratory Cycle, multiplied by 100.) Preliminary evidence indicates that the CQ provides a fairly good metric of vocal fold adduction (Titze, 1990; Scherer et al., 1995). Because of the well-recognized inability to identify the exact moment of vocal fold contact or loss of contact, a variable zero cross of the interpolated 25% peak-to-peak EGG amplitude level was used to delimit the EGG ''contact phase.'' The 25% crossing criterion was heuristically determined (Orlikoff, 1991, 1995). This baseline was found to represent the lowest level that would ensure freeing the data from waveform artifacts (Baken, 1987, 1992; Colton and Conture, 1990; Titze, 1990; Orlikoff, in press 1998) often associated with the EGG ''minimal-contact phase.'' Although different criterion levels (ranging roughly between 25% and 50% of maximal fold contact) have been used by investigators to derive %CQ, neither Rothenberg and

Mahshie (1988) nor Howard and his colleagues (1990, 1995) have shown that this substantially alters the measure or prevents comparison of different data sets, as long as the criterion used is strictly adhered to for each set. Although the computer demarcated EGG phases (A–C in Fig. 1) on the basis of digitized EGG maxima and minima, each EGG cycle analyzed was displayed on the computer screen along with the computer-selected demarcations to permit the experimenter to check the validity of measurement.

For the typical electroglottogram, the contact phase is asymmetrical; that is, the interval of increasing contact is shorter than that of decreasing contact. The degree of contact asymmetry is thought to reflect vocal fold tonus and to be particularly sensitive to vertical mucosal dynamics (Titze and Talkin, 1981; Childers et al., 1986; Titze, 1990). Orlikoff (1991) defined a dimensionless ratio to describe the symmetry of the EGG contacting and decontacting segments called the contact index (CI). It is defined as the ratio of the durational difference between these segments to the duration of the entire contact phase. Thus a CI that approaches −1 represents an EGG with a negligible increasing-contact duration, a CI of zero represents a perfectly symmetrical contact phase, and a CI that approaches +1 represents a waveform with a negligible decreasing-contact duration. The CI is inversely related to vocal intensity (Orlikoff, 1991), but can be expected to vary also with voice register (Fourcin, 1981; Kitzing, 1982; Roubeau et al., 1987; Schutte and Seidner, 1988; Welch et al., 1989) or the specific mode of vocal fold vibration (Hertegård and Gauffin, 1995).

Another measure, the relative EGG rise time (%RT), is the duration of the increasing contact segment relative to the entire vibratory cycle (Houben et al., 1992) and is tied to the closure event, which is crucial to the acoustic excitation of the vocal tract (Wechsler, 1977; Kakita, 1988; Fisher et al., 1995).

Lastly, as an index of the relative contact variability, the digitized maxima and minima of the EGG waveform were used to derive the percent amplitude perturbation. EGG amplitude perturbation was defined as the average absolute cycle-to-cycle difference in EGG amplitude, divided by the average wave amplitude, expressed in percent. For normal voice production EGG amplitude perturbation is typically about half the acoustic amplitude perturbation (shimmer) measured from a microphone signal (Linders et al., 1995; Orlikoff, 1991, 1995). Clearly, EGG and acoustic measures of amplitude perturbation are not comparable since the former is tied to vocal physiology while the latter is a metric of the acoustic yield. There is some preliminary evidence, however, that EGG amplitude perturbation, even if its relationship to voice acoustics is yet unclear, may provide a more sensitive gauge of vocal stability and function than acoustic perturbation measures (Haji et al., 1986; Horiguchi et al., 1987).

Data were analyzed using a repeated measures analysis of variance (ANOVA) model (CSS Ver. 2.1, StatSoft, Inc.). Subject sex and airflow direction (phonatory condition) were the across- and within-subject variables, respectively.

TABLE I. Mean fundamental frequency and short-term frequency perturbation data (standard deviations in parentheses) used by the male and female subjects comparing expiratory with inspiratory phonation. Statistical significance refers to effect of phonatory (EP versus IP) condition (**$p \leq 0.01$).

Measure	Expiratory phonation		Inspiratory phonation	
	Men	Women	Men	Women
F_0 (Hz)	101.86	202.84	133.55**	289.21**
	(15.55)	(31.63)	(25.32)	(67.83)
Jitter (%)	0.52	0.48	2.02**	1.84**
	(0.25)	(0.25)	(1.20)	(1.16)

II. RESULTS AND DISCUSSION

A. Frequency measures

Mean vocal fundamental frequency and percent jitter data are summarized in Table I. The mean expiratory F_0's of 101.86 and 202.84 Hz for the male and female subjects, respectively, are consistent with those reported in previous studies of phonation sustained at a subject-selected ''comfortable'' pitch (Brown et al., 1976; Murry et al., 1995; Orlikoff, 1990). For every subject tested the mean F_0 increased for inspiratory voice productions, rising on average to 133.55 Hz for the men and 289.21 Hz for the women. In general, each subject was quite consistent with respect to the frequency level used during expiratory and inspiratory samples. Recently, Grau et al. (1995) studied 20 normal full-term infants 2 days postpartum and likewise reported that the mean frequency of inspiratory cries were significantly higher than expiratory cries that immediately preceded and followed. Voice therapists working with adults have also recognized the tendency toward a rise in vocal pitch during inspiration (Boone and McFarlane, 1988) and have named exercises in which the patient—much as in the present experiment— alternates expiratory and inspiratory voice ''donkey breathing'' (Shulman, 1993) apparently because the two-toned vocalizations mimic the ''hee haw'' braying of the animal. The largest frequency shifts between expiratory and inspiratory conditions were 11.3 semitones for one female subject and 8.0 semitones for one of the men. On average, the men's phonations increased by 4.6 (s.d.=2.9) semitones, while the women's phonations increased by 5.9 (s.d.=2.7) semitones.

The relative frequency change shown by the male and female subjects was not significantly different ($p = 0.38$).

As with the mean F_0 associated with the expiratory voice samples, the men's and women's mean percent jitter of 0.52% and 0.48%, respectively, are consistent with values reported for healthy adults of similar age (Orlikoff, 1990, 1995; Ramig and Ringel, 1983; Wilcox and Horii, 1980). Despite high variability, the mean jitter of inspiratory segments was significantly higher [$F(1,28) = 22.68$, $p < 0.001$] than the expiratory samples, averaging 2.02% for the men and 1.84% for the women, with no significant jitter difference between the two sexes ($p = 0.71$).

B. Electroglottographic data

EGG data are summarized in Table II. The similar mean %CQs of 53.9% and 53.2% between the men and women, respectively, is consistent with the few published reports of subjects recorded at moderate pitches and loudnesses (Lindsey et al., 1988; Howard, 1995). The %CQ data also accord well with values provided by Higgins and Saxman (1991) when their inverted EGG waveforms are plotted as in the present experiment and the reported average minimum-contact duty cycles are subtracted from 100. Unlike the female subjects, whose inspiratory %CQs were essentially the same as those measured during normal EP, the men's inspiratory %CQs were significantly smaller [$FF(1,16) = 5.43$, $p < 0.05$] than they were during normal expiratory voice production. The reason for the difference between the

TABLE II. Means (and standard deviations) associated with the EGG contact quotient (%CQ), contact index (CI), rise time (%RT), and amplitude perturbation of the electroglottogram, comparing expiratory with inspiratory phonation. Statistical significance refers to effect of phonatory (EF versus IP) condition (*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$).

Measure	Expiratory phonation		Inspiratory phonation	
	Men	Women	Men	Women
%CQ	53.9	53.2	46.9*	54.6
	(4.4)	(5.0)	(3.8)	(7.1)
Cl	−0.62	−0.54	−0.39***	−0.24**
	(0.05)	(0.05)	(0.12)	(0.22)
%RT	9.6	12.2	14.2***	20.4**
	(1.1)	(0.7)	(3.0)	(5.1)
EGG amplitude perturbation (%)	1.18	1.51	2.36*	2.81*
	(0.88)	(0.85)	(0.99)	(1.30)
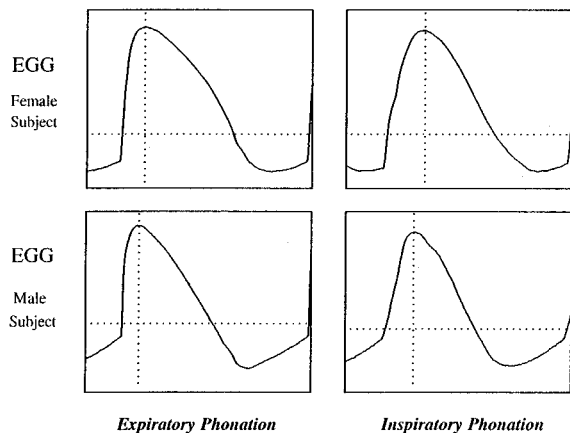
FIG. 2. Representative electroglottograms obtained during EP and IP for a 27-year-old female subject (above) and a 37-year-old male subject (below). The dotted vertical and horizontal lines represent the peak vocal fold contact area and the 25% baseline criterion, respectively.

TABLE III. Means (and standard deviations) associated with the phonatory airflow used by subjects during expiratory and inspiratory voice production.

	Airflow (mL/s)	
	Expiratory phonation	Inspiratory phonation
Men	158.5	−212.0
	(23.3)	(43.6)
Women	139.8	−228.0
	(30.2)	(55.2)
All subjects	149.2	−220.0
	(27.8)	(48.7)

sex groups with respect to the inspiratory condition is unclear and is a subject for further investigation.

A clear expiratory–inspiratory difference is evident from the contact index, CI. Statistical analysis showed both a sex $[F(1,28)=6.53, p<0.05]$ and phonatory condition $[F(1,28)=33.05, p<0.001]$ effect for CI. For the men, the mean CI of -0.62 for EP indicates that the increasing-contact interval represented approximately 19% of the contact phase on average, whereas during IP, the mean CI of -0.39 indicates that the increasing-contact interval accounted for about 30% of the entire contact phase. Among the female subjects, the mean CIs of -0.54 and -0.24 for expiratory and inspiratory efforts, respectively, showed a similar trend whereby the increasing-contact interval grew from representing about 23% to 38% of the contact phase. Comparing this same increasing-contact interval to the entire vibratory cycle, the %RT data likewise showed a significant phonatory condition effect $[F(1,28)=32.15, p<0.001]$, despite the reduced relative contact phase duration shown by the male subjects during IP. There was no significant sex effect ($p=0.26$) with respect to the %RT data. The average %RT values for EP correspond well to the rise time ratios reported by Houben *et al.* (1992) for the two men and one woman they tested.

The EGG characteristics described above are evident in some representative waveforms from one of the male and one of the female subjects shown in Fig. 2. Note the increased rise time and contact symmetry in both inspiratory samples. According to Titze's (1990) model of the physiologic bases of EGG changes, the differences shown in the inspiratory waveforms would be of the sort associated with, among other things, reduced glottal convergence. The ''thin'' quality often associated with IP (Miller and Schutte, 1990, 1993) thus does not appear necessarily to be due to a reduced contact duration, but rather to slower vocal fold approximation. The relatively quick sealing of the glottis and low open quotient described by Timcke *et al.* (1959) are not apparent in the present findings. This discrepancy may, in part, be due to the fact that data derived from their essentially two-dimensional glottographic width functions and the three-dimensional EGG are not fully equivalent. This has been well-recognized in attempts to derive open quotient data from the EGG.

Consistent with results reported by Horiguchi *et al.* (1987) and Orlikoff (1995), EGG amplitude perturbation measured during EP was quite similar for the male and female subjects, with a mean of 1.34% (s.d.=0.86). The percent EGG amplitude perturbation roughly doubled during IP, averaging 2.59% (s.d.=1.14) across subjects. This increased short-term EGG amplitude variation was statistically significant $[F(1,28)=11.85, p<0.005]$ and there was no significant sex effect ($p=0.29$). The greater instability of the degree of vocal fold contact during inspiratory voice production is consistent with the observed increase in frequency perturbation described above.

C. Airflow

Average phonatory airflow rates are shown in Table III. In general, the flows measured during EP were consistent with rates that have been reported for normal men and women (Biever and Bless, 1989; Holmberg *et al.*, 1988). The absolute airflow during IP was significantly higher than those used for expiratory voice production $[F(1,28)=25.12, p<0.001]$. On average the men's absolute airflow during vocal inhalation was 35.9% (s.d.=33.4) greater than during EP. The absolute airflow difference was greater for the women, whose mean inspiratory flow was 61.0% (s.d.=73.7) greater than during the expiratory task. However, the difference between the sexes was not statistically significant ($p=0.89$).

D. Videostroboscopic data

For the two men and two women examined by rigid stroboscopy, IPs were characterized by approximation of the upper margins of the vocal folds that preceded that of the lower margins, the reverse of what was observed during EP. Inspiratory voice production was also associated with thinner and more elongated vocal folds, on average represented by roughly a 14% increase from their length during EP, which is consistent with the observation of higher vocal F_0's associated with voice produced during inhalation.

In addition to a general widening of the supraglottis, the larynx typically assumed a lower neck position during inspiratory voice. With regard to quiet, especially deep, respiration, several investigators have observed that the larynx moves caudally during inspiration and rostrally during expi-

ration (Macklin, 1925; Mitchinson and Yoffey, 1947; Fink *et al.*, 1956). Fink (1974a) ascribed the inspiratory descent of the larynx to the lowering of the trachea which then acts to stretch and tow the laryngeal soft tissues. Additionally, at least in cats, EMG activity has been observed during inhalation in those muscles that act to lower the thyroid cartilage or to elevate the hyoid bone (Murakami and Kirchner, 1972, 1974). It has been hypothesized that increasing the distance between the hyoid and larynx would serve to flatten or ''unfold'' the aryepiglottic, ventricular, and vocal folds (Fink, 1974a,b, 1975), thus altering laryngeal biomechanics (Zenker and Glaninger, 1959; Shipp, 1975; Hoit *et al.*, 1993; Vilkman *et al.*, 1989, 1996; Iwarsson *et al.*, 1996). This action would also widen the laryngeal vestibule which has already been documented to occur during IP (Powers *et al.*, 1964; Lehmann, 1965).

During inspiratory voice segments, the quadrangular membranes appeared to be lengthened and to become taut as the vocal folds are displaced caudally. This laryngeal unfolding would increase the mucosal tension along the length of the laryngeal lumen and would be consistent with the relatively higher F_0's and more symmetrical EGG contact phase that are characteristic of ingressive voice. Inspiratory vocal fold vibration is also likely to be associated with a dramatic reduction or elimination of the upward elliptical deflection (mucosal upheaval) of the folds (Kumakawa *et al.*, 1990) that is a salient component of the normal expiratory oscillatory cycle (Matsushita, 1975; Baer, 1981; Saito *et al.*, 1981; Fukuda *et al.*, 1983).

E. Modes of vibration

Fukuda *et al.* (1987), using a mechanical vibrator applied to the thyroid prominence, drove laryngeal soft tissue longitudinally but nonetheless excited strong transverse vibratory modes in the vocal folds, similar to those seen during normal aerodynamically driven phonation. Given the apparent robustness of this mechanical coupling, it is perhaps not too surprising that the present subjects were able to maintain reasonable phonation when it was governed by an ingressive airflow. In fact, three of the men and one of the women were able to complete the present experiment using a low-frequency pulse (or fry) voice register and five of the men and six of the women were able to alternate inspiratory and expiratory phonation in a high-frequency loft (or falsetto) register (Fig. 3). This would suggest that at least a few of the several possible modes of vibration (Stevens, 1977; Berry and Titze, 1996; Orlikoff and Kahane, 1996) are not limited by the direction of transglottal airflow. Also, given that IP appears to be associated with the abduction of the ventricular folds and with ventricular distention, it would seem that pulse register phonation may not necessarily depend on a mass loading of vocal folds by the impinging ventricular folds as has been thought (Allen and Hollien, 1973).

III. GENERAL DISCUSSION

Inspiratory voice production is a volitional activity that can be made to vary with respect to the frequency and mode of vocal fold vibration. It appears to be associated with dis-

a. Pulse (fry) voice register
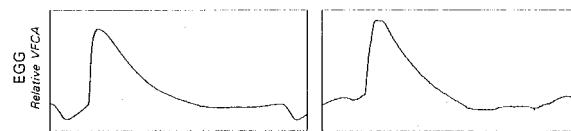
b. Loft (falsetto) voice register

Expiratory Phonation   *Inspiratory Phonation*
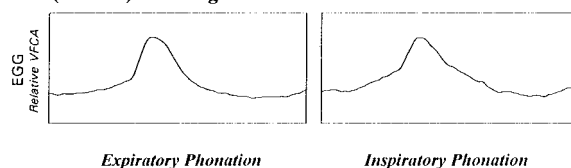
FIG. 3. Representative electroglottograms obtained during EP and IP at different voice registers [VFCA=vocal fold contact area]. Above a 31-year-old male subject alternated between egressive and ingressive airflow using a pulse register (mean EP and IP F_0's=36.1 and 48.9 Hz, respectively). Below a 23-year-old female subject alternated flow direction using a loft register voice (mean EP and IP F_0's=433.1 and 435.6 Hz, respectively).

tention of the laryngeal ventricles and lengthening of the quadrangular membrane. In all of these respects it differs from the abnormal sound production of inspiratory stridor, which is often associated with supraglottic lesions (Finucane and Santora, 1988), and of wheezing, which is usually due to subglottal constriction or obstruction and is more pronounced during exhalation (Forgacs, 1967; Williams and Phelan, 1975; Gavriely *et al.*, 1984). Although IPs elicited in the present study were typically associated with greater short-term acoustic and electroglottographic variability, they could not be considered ''nonharmonic'' as previously described by Miller and Schutte (1990, 1993), although the reportedly ''tight adductory force'' used by their subjects may have been of some consequence in this regard.

The present study represents a first examination of the acoustic and electroglottographic characteristics of IP. More detailed and directed study of inspiratory voice production and the associated acoustic product is needed. In particular, a better understanding of inspiratory laryngeal biomechanics may explain why IP may be easier or more efficient for some to produce (Till *et al.*, 1983) and why certain voice patients are constrained or choose to speak exclusively on ingressive air (e.g., Frint and Kelemen, 1969). Future investigation should address how the articulatory function of the larynx is affected by laryngeal depression and by an ingressive airflow.

Speech and vocalization using an ingressive airflow provides an interesting window on phonatory dynamics that may help to clarify certain aspects of normal vocal neurology, laryngeal mechanics, and aerodynamics. Beyond this, if IP is found to be a useful technique for vocal rehabilitation, its physiological characteristics must be understood. At present our relative ignorance limits its applicability in the assessment and management of disorder.

Terme, Italy, and on February 5, 1997 at the Midwinter Meeting of the Association for Research in Otolaryngology (ARO) in St. Petersburg, FL.

Allen, E. L., and Hollien, H. (**1973**). ''A laminagraphic study of pulse (vocal fry) register phonation,'' Folia Phoniatr. **25**, 241–250.

Baer, T. (**1981**). ''Observation of vocal fold vibration: Measurement of excised larynges,'' in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo), pp. 119–133.

Baken, R. J. (**1987**). *Clinical Measurement of Speech and Voice* (College-Hill, San Diego, CA).

Baken, R. J. (**1992**). ''Electroglottography,'' J. Voice **6**, 98–110.

Berry, D. A., and Titze, I. R. (**1996**). '' Normal modes in a continuum model of vocal fold tissues,'' J. Acoust. Soc. Am. **100**, 3345–3354.

Biever, D. M., and Bless, D. M. (**1989**). ''Vibratory characteristics of the vocal folds in young adult and geriatric women,'' J. Voice **3**, 120–131.

Boone, D. R., and McFarlane, S. C. (**1988**). *The Voice and Voice Therapy* (Prentice-Hall, Englewood Cliffs, NJ).

Brown, W. S., Jr., Murry, T., and Hughes, D. (**1976**). ''Comfortable effort level: An experimental variable,'' J. Acoust. Soc. Am. **60**, 696–699.

Bunn, J. C., and Mead, J. (**1971**). ''Control of ventilation during speech,'' J. Appl. Physiol. **31**, 870–872.

Childers, D. G., Hicks, D. M., Moore, G. P., and Alsaka, Y. A. (**1986**). ''A model of vocal fold vibratory motion, contact area, and the electroglottogram,'' J. Acoust. Soc. Am. **80**, 1309–1320.

Colton, R. H., and Conture, E. G. (**1990**). ''Problems and pitfalls of electroglottography,'' J. Voice **4**, 10–24.

D'Antonio, L., Lotz, W., Chait, D., and Netsell, R. (**1987**). ''Perceptual-physiologic approach to evaluation and treatment of dysphonia,'' Ann. Otol. Rhinol. Laryngol. **96**, 187–190.

Davis, P. J., Bartlett, D., Jr., and Luschei, E. S. (**1993**). ''Coordination of the respiratory and laryngeal systems in breathing and vocalization,'' in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. R. Titze (Singular, San Diego, CA), pp. 189–209.

Faure, M.-A., and Muller, A. (**1992**). ''Stroboscopy,'' J. Voice **6**, 139–148.

Fink, B. R. (**1974a**). ''Folding mechanism of the human larynx,'' Acta Oto-Laryngol. **78**, 124–128.

Fink, B. R. (**1974b**). ''Spring mechanisms in the human larynx,'' Acta Oto-Laryngol. **78**, 295–304.

Fink, B. R. (**1975**). *The Human Larynx: A Functional Study* (Raven, New York).

Fink, B. R., Basek, M., and Epanchin, V. (**1956**). ''The mechanisms of opening of the human larynx,'' Laryngoscope **66**, 410–425.

Finucane, B. T., and Santora, A. H. (**1988**). *Principles of Airway Management* (F. A. Davis, Philadelphia, PA), p. 76.

Fisher, K., Scherer, R., Swank, P., Giddens, C., and Patten, D. (**1995**). ''Electroglottographic tracking of phonatory response to Botox,'' NCVS Stat. Prog. Rep. **8**, 97–111.

Forgacs, P. (**1967**). ''Crackles and wheezes,'' Lancet **2**, 203–205.

Frint, T., and Kelemen, A. (**1969**). ''Inspiratorische Stimmbildung psychogenen Ursprunges,'' Folia Phoniatr. **21**, 105–111.

Fourcin, A. J. (**1981**). ''Laryngographic assessment of phonatory function,'' ASHA Rep. **11**, 116–127.

Fukuda, H., Saito, S., Kitahara, S., Isogai, Y., Makino, K., Tsuzuki, T., Kogawa, N., and Ono, H. (**1983**). ''Vocal fold vibration in excised larynges viewed with an x-ray stroboscope and an ultra-high-speed camera,'' in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego, CA), pp. 238–252.

Fukuda, H., Muta, H., Kanou, S., Takayama, E., Fujioka, T., Kawaida, T., Tatehara, T., and Saito, S. (**1987**). ''Response of vocal folds to externally induced vibrations: Basic study and its clinical application,'' in *Laryngeal Function in Phonation and Respiration*, edited by T. Baer, C. T. Sasaki, and K. S. Harris (College-Hill, Boston), pp. 366–377.

Gauffin, J., Binh, N., Ananthapadmanabha, T. V., and Fant, G. (**1983**). ''Glottal geometry and volume velocity waveform,'' in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego, CA), pp. 194–201.

Gauffin, J., and Liljencrants, J. (**1988**). ''Modelling the airflow in the glottis,'' Ann. Bull. Res. Inst. Logoped. Phoniatr. **22**, 39–50.

Gavriely, N., Palty, Y., Alroy, G., and Grotberg, J. B. (**1984**). ''Measurement and theory of wheezing breath sounds,'' J. Appl. Physiol. **57**, 481–492.

Grau, S. M., Robb, M. P., and Cacace, A. T. (**1995**). ''Acoustic correlates of inspiratory phonation during infant cry,'' J. Speech Hear. Res. **38**, 373–381.

Haji, T., Horiguchi, S., Baer, T., and Gould, W. J. (**1986**). ''Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation,'' J. Acoust. Soc. Am. **80**, 58–62.

Hertegård, S., and Gauffin, J. (**1995**). ''Glottal area and vibratory patterns studied with simultaneous stroboscopy, flow glottography, and electroglottography,'' J. Speech Hear. Res. **38**, 85–100.

Higgins, M. B., and Saxman, J. H. (**1991**). ''A comparison of selected phonatory behaviors of healthy aged and young adults,'' J. Speech Hear. Res. **34**, 1000–1010.

Hirano, M., and Bless, D. M. (**1993**). *Videostroboscopic Examination of the Larynx* (Singular, San Diego, CA).

Hoit, J. D., Solomon, N. P., and Hixon, T. J. (**1993**). ''Effect of lung volume on voice onset time (VOT),'' J. Speech Hear. Res. **36**, 516–521.

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.

Horiguchi, S., Haji, T., Baer, T., and Gould, W. J. (**1987**). ''Comparison of electroglottographic and acoustic waveform perturbation measures,'' in *Laryngeal Function in Phonation and Respiration*, edited by T. Baer, C. T. Sasaki, and K. S. Harris (College-Hill, Boston), pp. 509–518.

Houben, G. B., Buekers, R., and Kingma, H. (**1992**). ''Characterization of the electroglottographic waveform: A primary study to investigate vocal fold functioning,'' Folia Phoniatr. **44**, 269–281.

Howard, D. M. (**1995**). ''Variation of electrolaryngraphically derived closed quotient for trained and untrained adult female singers,'' J. Voice **9**, 163–172.

Howard, D. M., Lindsey, G. A., and Allen, B. (**1990**). ''Towards the quantification of vocal efficiency,'' J. Voice **4**, 205–212.

Iwarsson, J., Thomasson, M., and Sundberg, J. (**1996**). ''Effects of lung volume on the glottal voice source,'' Karolinska Inst. Phoniatr. Logoped. Prog. Rep. **10**, 9–18.

Kakita, Y. (**1988**). ''Simultaneous observation of the vibratory pattern, sound pressure, and airflow signals using a physical model of the vocal folds,'' in *Vocal Physiology: Voice Production, Mechanisms and Functions*, edited by O. Fujimura (Raven, New York), pp. 207–218.

Kitzing, P. (**1982**). ''Photo- and electroglottographic recording of the laryngeal vibratory pattern during different registers,'' Folia Phoniatr. **34**, 234–241.

Kollár, A. (**1973**). ''Inspiratorische Phonation bei der Behandlung von Stimmstorungen,'' Folia Phoniatr. **25**, 221–224.

Kollár, A. (**1989**). ''Príspevek k neprímé laryngoskopii [Contribution to indirect laryngoscopy],'' Ceskoslov. Otorhinolaryngol. **38**, 114–115.

Kumakawa, K., Yamaguchi, H., and Hirose, H. (**1990**). ''Image analysis of the contour of paralyzed vocal cord during phonation,'' J. Oto-Rhino-Laryngol. Soc. Jpn. **93**, 61–70.

Lehmann, Q. H. (**1965**). ''Reverse phonation: A new maneuver for examining the larynx,'' Radiology **84**, 215–222.

Linders, B., Massa, G. G., Boersma, B., and Dejonckere, P. H. (**1995**). ''Fundamental voice frequency and jitter in girls and boys measured with electroglottography: influence of age and height,'' Int. J. Ped. Otorhinolaryngol. **33**, 61–65.

Lindsey, G., Breen, A. P., and Fourcin, A. J. (**1988**). ''Glottal closed time as a function of prosody, style and sex in English,'' in *Proceedings of the 7th Federation of Acoustical Societies of Europe Symposium* (Edinburgh) **3**, 22–26.

Macklin, C. (**1925**). ''X-ray studies on bronchial movements,'' Am. J. Anat. **35**, 303–320.

Matsushita, H. (**1975**). ''The vibratory mode of the vocal folds in the excised larynx,'' Folia Phoniatr. **27**, 7–18.

Mead, J., and Bunn, J. C. (**1974**). ''Speech as breathing,'' in *Ventilatory and Phonatory Control Systems*, edited by B. Wyke (Oxford, New York), pp. 33–38.

Miller, D. G., and Schutte, H. K. (**1990**). ''Feedback from spectrum analysis applied to the singing voice,'' J. Voice **4**, 329–334.

Miller, D. G., and Schutte, H. K. (**1993**). ''Physical definition of the 'Flageolet régister','' J. Voice **7**, 206–212.

Mitchinson, A., and Yoffey, J. (**1947**). ''Respiratory displacement of larynx, hyoid bone, and tongue,'' J. Anat. **81**, 118–120.

Murakami, Y., and Kirchner, J. A. (**1972**). ''Respiratory movements of the vocal cords,'' Laryngoscope **82**, 454–467.

Murakami, Y., and Kirchner, J. A. (**1974**). ''Respiratory activity of the

external muscles: An electromyographic study of the cat,'' in *Ventilatory and Phonatory Control Systems*, edited by B. Wyke (Oxford, New York), pp. 430–448.

Murry, T., Brown, W. S., Jr., and Morris, R. J. (**1995**). ''Patterns of fundamental frequency for three types of voice samples,'' J. Voice **9**, 282–289.

Orlikoff, R. F. (**1990**). ''The relationship of age and cardiovascular health to certain acoustic characteristics of male voices,'' J. Speech Hear. Res. **33**, 450–457.

Orlikoff, R. F. (**1991**). ''Assessment of the dynamics of vocal fold contact from the electroglottogram: Data from normal male subjects,'' J. Speech Hear. Res. **34**, 1066–1072.

Orlikoff, R. F. (**1995**). ''Vocal stability and vocal tract configuration: An acoustic and electroglottographic investigation,'' J. Voice **9**, 173–181.

Orlikoff, R. F. (**in press**). ''Scrambled EGG: The uses and abuses of electroglottography,'' Phonoscope.

Orlikoff, R. F., and Kahane, J. C. (**1996**). ''Laryngeal structure and function,'' in *Principles of Experimental Phonetics*, edited by N. J. Lass (Mosby, St. Louis, MO), pp. 112–181.

Phillipson, E. A., McClean, P. A., Sullivan, C. E., and Zamel, N. (**1978**). ''Interaction of metabolic and behavioral respiratory control during hypercapnia and speech,'' Am. Rev. Respir. Dis. **117**, 903–909.

Powers, W. E., Holtz, S., and Ogura, J. (**1964**). ''Contrast examination of the larynx and pharynx: Inspiratory phonation,'' Am. J. Roentgenol. Rad. Ther. Nucl. Med. **92**, 40–42.

Ramig, L. A., and Ringel, R. L. (**1983**). ''Effects of physiological aging on selected acoustic characteristics of voice,'' J. Speech Hear. Res. **26**, 22–30.

Reich, A., and Till, J. (**1983**). ''Phonatory and manual reaction times of women with idiopathic spasmodic dysphonia,'' J. Speech Hear. Res. **26**, 10–18.

Reich, A., Till, J., and Goldsmith, H. (**1981**). ''Laryngeal and manual reaction times of stuttering and nonstuttering adults,'' J. Speech Hear. Res. **24**, 192–196.

Rothenberg, M., and Mahshie, J. J. (**1988**). ''Monitoring vocal fold abduction through vocal fold contact area,'' J. Speech Hear. Res. **31**, 338–351.

Roubeau, C., Chevrie-Muller, C., and Arabia-Guidet, C. (**1987**). ''Electroglottographic study of the changes of voice registers,'' Folia Phoniatr. **39**, 280–289.

Saito, S., Fukuda, H., Isogai, Y., and Ono, H. (**1981**). ''X-ray stroboscopy,'' in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo, Tokyo), pp. 95–106.

Scherer, R. C., Gould, W. J., Titze, I. R., Meyers, A. D., and Sataloff, R. T. (**1988**). ''Preliminary evaluation of selected acoustic and glottographic measures for clinical phonatory function analysis,'' J. Voice **2**, 230–244.

Scherer, R. C., and Guo, C.-G. (**1991**). ''Generalized translaryngeal pressure coefficients for a wide range of laryngeal configurations,'' in *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*, edited by J. Gauffin and B. Hammarberg (Singular, San Diego, CA), pp. 83–90.

Scherer, R. C., and Titze, I. R. (**1983**). ''Pressure-flow relationships in a model of the laryngeal airway with a diverging glottis,'' in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego, CA), pp. 179–193.

Scherer, R. C., and Titze, I. R. (**1987**). ''The abduction quotient related to vocal quality,'' J. Voice **1**, 246–251.

Scherer, R. C., Vail, V. J., and Rockwell, B. (**1995**). ''Examination of the laryngeal adduction measure EGGW,'' in *Producing Speech: Contemporary Issues for Katherine Stafford Harris*, edited by F. Bell-Berti and L. J. Raphael (AIP, New York), pp. 269–289.

Schutte, H. K., and Seidner, W. W. (**1988**). ''Registerabhängige Differenzierung von Elektroglottogrammen,'' Sprache Stimme Gehör. **12**, 59–62.

Shipp, T. (**1975**). ''Vertical laryngeal position during continuous and discrete vocal frequency change,'' J. Speech Hear. Res. **18**, 707–718.

Shulman, S. (**1993**). ''Symptom modification for abductor spasmodic dysphonia,'' in *Voice Therapy: Clinical Studies*, edited by J. C. Stemple (Mosby, St. Louis, MO), pp. 128–131.

Stevens, K. N. (**1977**). ''Physics of laryngeal behavior and larynx modes,'' Phonetica **34**, 264–279.

Sullivan, W. J., Peters, G. M., and Enright, P. L. (**1984**). ''Pneumotachographs: Theory and clinical application,'' Respiratory Care **29**, 736–749.

Till, J. A., Reich, A., Dickey, S., and Seiber, J. (**1983**). ''Phonatory and manual reaction times of stuttering and nonstuttering children,'' J. Speech Hear. Res. **26**, 171–180.

Timcke, R., von Leden, H., and Moore, P. (**1958**). ''Laryngeal vibrations: Measurements of the glottic wave. Part I—The normal vibratory cycle,'' Acta Oto-Laryngol. **68**, 1–19.

Timcke, R., von Leden, H., and Moore, P. (**1959**). ''Laryngeal vibrations: Measurements of the glottic wave. Part II—Physiologic variations,'' Acta Oto-Laryngol. **69**, 438–444.

Titze, I. R. (**1984**). ''Parameterization of the glottal area, glottal flow, and vocal fold contact area,'' J. Acoust. Soc. Am. **75**, 570–580.

Titze, I. R. (**1989**). ''A four parameter model of the glottis and vocal fold contact area,'' Speech Commun. **8**, 191–201.

Titze, I. R. (**1990**). ''Interpretation of the electroglottographic signal,'' J. Voice **4**, 1–9.

Titze, I. R., and Talkin, D. (**1981**). ''Simulation and interpretation of glottographic waveforms,'' ASHA Rep. **11**, 48–55.

Vieira, M. N., McInnes, F. R., and Jack, M. A. (**1996**). ''Analysis of the effects of electroglottographic baseline fluctuation on the F_0 estimation in pathological voices,'' J. Acoust. Soc. Am. **99**, 3171–3178.

Vieira, M. N., McInnes, F. R., and Jack, M. A. (**1997**). ''Comparative assessment of electroglottographic and acoustic measures of jitter in pathological voices,'' J. Speech Hear. Res. **40**, 170–182.

Vilkman, E., Sonninen, A., Hurme, P., and Körkkö, P. (**1996**). ''External laryngeal frame function in voice production revisited: A review,'' J. Voice **10**, 78–92.

Vilkman, E., Aaltonen, O., Raimo, I., Arajärvi, P., and Oksanen, H. (**1989**). ''Articulatory hyoid-laryngeal changes vs. cricothyroid muscle activity in the control of intrinsic F_0 of vowels,'' J. Phonet. **17**, 193–203.

Watterson, T. L., and McFarlane, S. C. (**1991**). ''Transoral and transnasal laryngeal endoscopy,'' Sem. Speech Lang. **12**, 77–87.

Wechsler, E. (**1977**). ''A laryngographic study of voice disorders,'' Br. J. Dis. Comm. **12**, 9–22.

Welch, G. F., Sergeant, D. C., and MacCurtain, F. (**1989**). ''Xeroradiographic-electrolaryngographic analysis of male vocal registers,'' J. Voice **3**, 244–256.

Wilcox, K. A., and Horii, Y. (**1980**). ''Age and changes in vocal jitter,'' J. Gerontol. **35**, 194–198.

Williams, H. E., and Phelan, P. D. (**1975**). *Respiratory Illness in Children* (Blackwell, Oxford), pp. 106–115.

Winkworth, A. L., Davis, P. J., Ellis, E., and Adams, R. D. (**1994**). ''Variability and consistency in speech breathing during reading: Lung volumes, speech intensity, and linguistic factors,'' J. Speech Hear. Res. **37**, 535–556.

Winkworth, A. L., Davis, P. J., Adams, R. D., and Ellis, E. (**1995**). ''Breathing patterns during spontaneous speech,'' J. Speech Hear. Res. **38**, 124–144.

Wyke, B. D. (**1983**). ''Neuromuscular control systems in voice production,'' in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego, CA), pp. 71–76.

Zenker, W., and Glaninger, J. (**1959**). ''Die Stärke des Trachealzuges beim lebenden Menschen und seine Bedeutung für die Kehlkopfmechanik,'' Z. Biol. **111**, 154–166.

Velopharyngeal function in young and older adult speakers: Evidence from aerodynamic studies

David J. Zajac[a]

Department of Dental Ecology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Intraoral air pressure (P_0), nasal airflow, and nasal air volume characteristics were determined from 63 adult speakers during production of oral and nasal consonants. The speakers were categorized into two age groups. Young speakers (21 males, 21 females) ranged in age from 18 to 37 years (mean=24); older speakers (10 males, 11 females) ranged in age from 68 to 83 years (mean=75). Speakers produced the syllables /pi/ and /si/ and the word ''hamper'' at self-determined loudness levels during reiterant speech tasks. Results indicated similar velopharyngeal function between young and older adults as reflected by nasal airflow and air volume measures. Older adults, however, exhibited significantly higher levels of P_0 than younger adults during production of /p/ in ''hamper.'' In addition, (a) male speakers exhibited higher P_0 than females during production of /p/ in ''hamper,'' (b) all speakers demonstrated declination of P_0 during production of /p/ in ''hamper'' as a function of word position in the utterance, and (c) male and female speakers exhibited different patterns of nasal air volume during production of ''hamper'' as a function of word position in the utterance. The intraoral air pressure results suggest differences in respiratory function between young and older speakers. The nasal air volume results suggest a gender specific velopharyngeal declination effect. Implications for velar motor control strategies are discussed. © *1997 Acoustical Society of America.* [S0001-4966(97)01009-6]

PACS numbers: 43.70.Aj, 43.70.Bk [AL]

INTRODUCTION

The effects of aging on velopharyngeal function during speech production have received relatively little attention. Hutchinson *et al.* (1978) reported several patterns of nasalance in older speakers which they interpreted as evidence of deterioration in velopharyngeal function. One pattern was characterized by pervasive elevation of nasalance levels during reading of an oral speech sample. Hutchinson *et al.* (1978) suggested that ''generalized neuromuscular weakness'' may have contributed to this finding. Another pattern was characterized by a progressive increase in nasalance as a function of speaking duration. Hutchinson *et al.* (1978) speculated that either fatigue and/or a decrease in levels of neurotransmitters may have contributed to this finding. Although statistical analyses of the associations among gender, age, and nasalance patterns were not reported by Hutchinson *et al.* (1978), the data suggested a possible age by sex effect. Specifically, of 31 speakers aged 66–88 yr, only one male but five females were categorized as exhibiting a pattern of progressive nasalance as a function of speaking duration. Of ten younger speakers, however, none exhibited this pattern (Table 3, p. 478). The possibility of an age by gender effect relative to nasalance should not be surprising given the fairly consistent body of literature which has indicated gender differences in various aspects of velopharyngeal function (e.g., Hoit *et al.*, 1994; Kuehn, 1976; McKerns and Bzoch, 1970; Seaver *et al.*, 1991; Thompson and Hixon, 1979; Zajac and Mayo, 1996).

Velopharyngeal function during speech production of young and older adults was investigated using an aerodynamic approach by Hoit *et al.* (1994). Eighty subjects, evenly distributed by gender, were categorized into four age groups: 20–30, 40–50, 60–70, and 80+ yr. Subjects produced both nasal and non-nasal phonemes in isolation, in reiterant syllables, and in syllables embedded within a carrier phrase. Hoit *et al.* (1994) reported no significant differences in nasal airflow among any of the age groups during production of non-nasal phonemes. The investigators explained the apparent discrepancy to the findings of Hutchinson *et al.* (1978) on the basis of differences in the types of measurements obtained. They speculated that older speakers may exhibit elevated nasalance even in the presence of adequate velopharyngeal closure due to an increased transfer of acoustic energy through palatal structures that may be less dense because of atrophy. Although this may explain the patterns of pervasive elevation of nasalance reported by Hutchinson *et al.* (1978), it does not explain the patterns of progressive nasalance. Indeed, because of the measurement procedures employed by Hoit *et al.* (1994), it is unlikely that progressive deterioration of velopharyngeal function within an utterance or phrase, if present, would have been detected. For example, measurements of nasal airflow were obtained at the midpoint of sustained productions and at the midpoint of the initial vowel in the syllables embedded in the carrier phrases. Although Hoit *et al.* (1994) obtained measurements from the third, fourth, and fifth vowels of the reiterant syllables, they did not present any statistical analyses of these sequential measurements.

Recently, a phenomenon which may help explain the

---

findings of progressive nasalance was reported in the literature. Krakow *et al.* (1995) suggested that velar function during an utterance was characterized by a kinematic ''declination effect,'' similar to that observed for other articulators. Specifically, Krakow *et al.* (1995) reported that velar height systematically decreased for two of three adult subjects as a function of utterance length. One of the speakers who exhibited declination was female; the sex of the other speakers was not specified. The investigators, however, suggested that velar declination may be part of a general phenomenon of speech production that affects all of the supralaryngeal articulators. For example, Vatikiotis-Bateson and Fowler (1988) reported declination effects involving the magnitude and peak velocity of labial movements during reiterant speech tasks for English speakers. Similarly, Vayra and Fowler (1992) demonstrated declination of $F1$ and $F2$ frequencies of vowels produced in multisyllabic pseudo words by Italian speakers. These investigators also reported declination in jaw opening of stressed vowels as a function of utterance length. If velar declination is a general phenomenon of speech production as suggested by Krakow *et al.* (1995), then this effect may have contributed to the progressive nasalance patterns reported by Hutchinson *et al.* (1978).

The purpose of the present study was to further explore velopharyngeal function in young and older adult speakers. The primary objective was to determine if older adults differ from younger adults relative to the aerodynamic integrity of the velopharyngeal structures during production of oral and nasal consonants. A secondary objective was to determine if older adults exhibit a velopharyngeal declination effect as reflected by nasal airflow measures during a reiterant speech task. Finally, because of the possible interaction of age and gender as suggested by previous studies, the sex of the speakers was systematically evaluated.

I. METHOD

A. Subjects

Sixty-three adult speakers participated in the study. They were categorized into two groups based upon age. Younger speakers consisted of 21 males and 21 females with a mean age of 24 yr (range=18–37). The younger speakers previously participated in a study of aerodynamic and temporal aspects of velopharyngeal function (Zajac and Mayo, 1996). The older speakers consisted of 10 males and 11 females with a mean age of 75 yr (range=68–83). All subjects were Caucasian, native speakers of standard American English. All subjects were in good health and free of upper respiratory infections at the time of testing. Fifteen of the older subjects were recruited from an ongoing study of sleep apnea. As controls, they completed a comprehensive health and psychiatric assessment prior to the study. The additional six older subjects reported no chronic medical problems. All subjects reported a negative history of overt or submucous cleft palate, hearing loss, and voice pathology. Smoking history of the subjects, however, was not obtained. Consistent with the self-reports, voice and resonance of all subjects were judged to be within normal limits for age.

B. Procedures

The pressure-flow method was used to detect intraoral air pressure, nasal air pressure, and nasal airflow during production of syllables and words (Warren *et al.*, 1985; Zajac and Mayo, 1996). A polyethylene catheter (1.67-mm internal diameter) was placed in the subject's mouth behind the upper incisors. A cork/catheter assembly was used to occlude the subject's least patent nostril. Patency of the nostrils was determined by a combination of perceptual evaluation, subjective report, and/or rhinomanometric testing. To detect intraoral and nasal air pressures, the two catheters were connected to the high ports of separate bidirectional pressure transducers referenced to atmosphere (Setra model 239). The pressure transducers were calibrated with a well-type manometer. To detect nasal airflow, a plastic tube was snugly fitted to the subject's most patent nostril and connected to a heated pneumotachograph (Fleisch #1) coupled to a differential pressure transducer (Setra model 239). The pneumotachograph and pressure transducer were calibrated with a captured air supply and rotameter. The frequency response of the pressure transducers was flat within 1% over the range of interest from 0 to 20 Hz. After the pressure and airflow catheters were comfortably in place, the subject repeated the syllable /pi/, the syllable /si/, and the word ''hamper.'' Each speech sample was produced five times during a single breath group with self-determined rate, loudness, and vocal quality. The pressure-flow data were low-pass filtered at 15 Hz and digitized to a computer at a rate of 200 samples/second with 12-bit resolution. PERCI-PSCOPE software, Version 1.1 (Microtronics, Inc.) was used to obtain the pressure-flow measurements.

C. Data analysis

Figure 1 illustrates a typical pressure-flow record for one of the productions of ''hamper.'' The volume rate of nasal airflow and intraoral air pressure (P_0) associated with the /m/ segment were determined for each subject at peak airflow (point D in Fig. 1); nasal airflow and P_0 associated with the /p/ segment were determined at peak pressure (point F). Mean values were calculated based upon the second, third, and fourth productions of the word. Air volume associated with the /m/ segment was determined by integration of the airflow signal. The beginning of the /m/ segment was identified when the initial rise in nasal airflow which preceded peak oral pressure exceeded 10 ml/s (point C in Fig. 1); the end of the segment was defined when nasal airflow decreased to 10 ms/s or less (point E). Justification for the use of a 10 ml/s criterion is given below. Two ratios of velopharyngeal function were also computed from the integrated airflow signal. One ratio compared the air volumes associated with points A–C to points A–E in Fig. 1. The segment A–C includes nasal airflow associated with the word boundary within the breath group in addition to nasal airflow associated with the /h/ and /æ/ segments. The segment A–E consists of nasal airflow associated with the entire gesture. A second ratio compared the nasal air volume associated with just the /æ/ segment (points B–C) to the /æ/ and /m/ segments (points B–E) within ''hamper.'' These ratios were de-
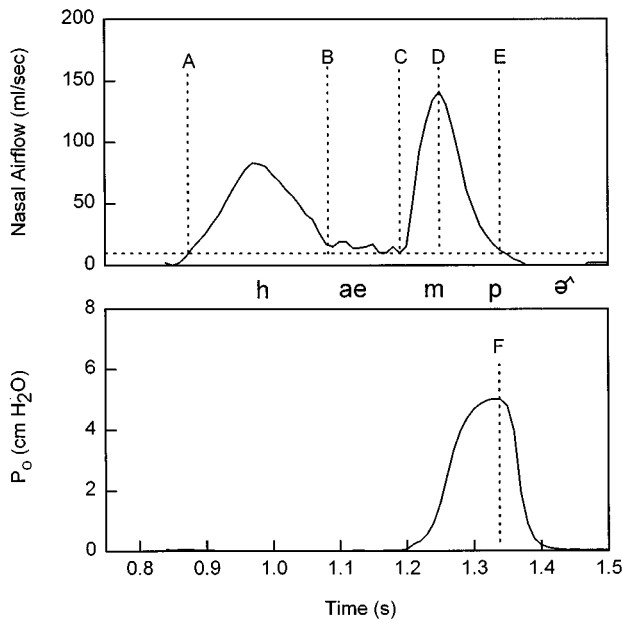
FIG. 1. Nasal airflow and intraoral air pressure (P_0) signals of a young female speaker during production of ''hamper.'' Dotted vertical lines define measurement points and segments as described in the text. The dotted horizontal line indicates the flow measurement criterion. Key: A=begin flow for word, B=begin flow for /æ/, C=begin flow for /m/, D=peak flow for /m/, E=end flow for word, F=peak pressure for /p/.

FIG. 2. Nasal airflow and audio signals of a young male speaker producing ''hamper.'' The /m/ segment as defined aerodynamically shows close correspondence to the acoustic manifestation of /m/.

termined because Krakow (1993) has suggested that ''a more complete model of speech motor organization would specify the various nonsegmental as well as segmental influences on the articulators'' (p. 88). A comparison of the two ratios, therefore, was anticipated to provide insight into the effects of word termination and initiation within a breath group (a nonsegmental influence) on velopharyngeal function. Because acoustic or kinematic data were not collected as part of the present study, it must be emphasized that the phonetic segments identified in Fig. 1 were defined by aerodynamic criteria. Additional voice and nasal airflow data, however, were obtained from one of the speakers and are illustrated in Fig. 2. These data suggest that the segments identified in Fig. 1 are consistent with the corresponding phonetic labels.[1]

During production of the syllables /pi/ and /si/, nasal airflow was determined at peak oral pressures. If peak nasal airflow occurred at locations other than the pressure peaks during production of either syllable, then these values were also recorded. Mean flow values were calculated based upon the second, third, and fourth repetitions of the syllables. As was done for the word ''hamper,'' the initial and final productions of the syllables were not analyzed to avoid nasal airflow artifacts due to velopharyngeal adjustments associated with the onset and offset of the breath group. Based upon the peak nasal airflow values from /pi/ and /si/, speakers were categorized as having either complete or incomplete velopharyngeal closure. Complete closure was defined as a volume rate of nasal airflow that was 10 ml/s or less. This criterion was selected because, as discussed by Thompson and Hixon (1979), positive or negative nasal airflow may occur within this range due to muscular contractions of the velum in the presence of airtight velopharyngeal closure.
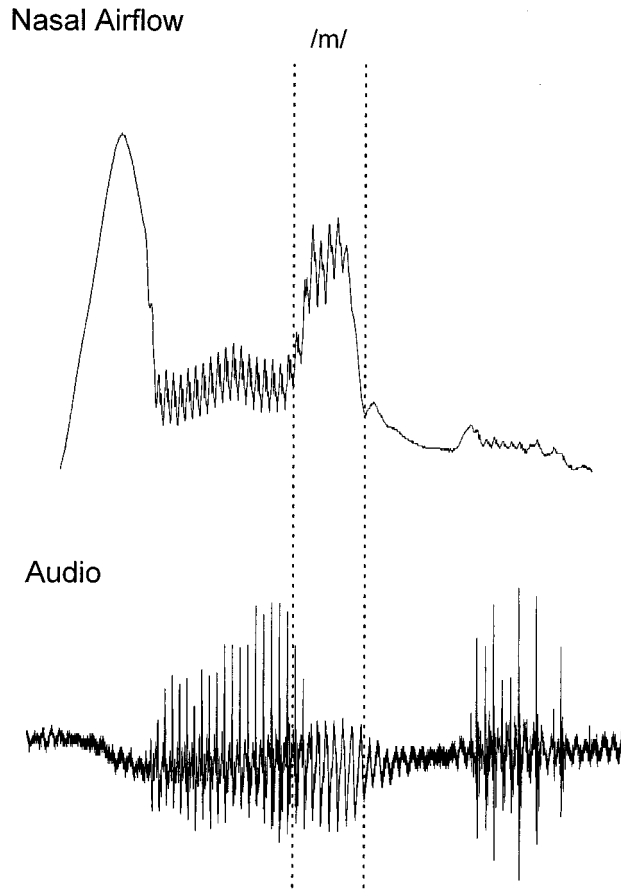
D. Statistical analyses

Descriptive statistics were calculated for the dependent variables as a function of age and gender. Analysis of variance (ANOVA) procedures with age and gender as factors were used to assess differences among P_0, nasal airflow, and nasal air volume measures associated with ''hamper.'' Chi squares were used to assess age and/or gender differences in the number of speakers categorized as having complete or incomplete velopharyngeal closure based upon the syllables /pi/ and /si/. Finally, velopharyngeal declination was assessed by performing repeated measures ANOVAs with age and gender as factors using the nasal air volume and P_0 data obtained from the second, third, and fourth repetitions of the word ''hamper.'' Probability levels were set at 0.05 for all statistical tests.

II. RESULTS

A. Age and gender effects

Means and standard deviations of P_0 and nasal airflow during production of /m/ and /p/ in ''hamper'' as a function of age and gender are presented in Table I. Results of the ANOVAs indicated no significant differences relative to P_0 associated with /m/. Relative to /p/, significant main effects of both gender ($F=14.5$, $df=1$, $p=0.0001$) and age ($F=8.1$, $df=1$, $p=0.006$) were indicated. Male speakers, re-

TABLE I. Means and standard deviations (s.d.) of intraoral air pressure (P_0) and nasal airflow associated with /m/ and /p/ in ''hamper'' as a function of age and gender.

| | Young speakers | | | | Older speakers | | | |
| | Male | | Female | | Male | | Female | |
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
P_0 (cm H_2O)								
/m/	1.5	0.9	1.1	0.5	1.3	0.3	1.2	0.3
/p/	5.9	1.6	4.8	1.0	7.2	1.4	5.5	1.2
Nasal airflow (ml/s)								
/m/	149	73	128	35	152	47	145	46
/p/	20	21	9	8	16	13	13	13

gardless of age, exhibited greater P_0 than females; older speakers, regardless of gender, exhibited greater P_0 than young speakers. Relative to nasal airflow, the ANOVAs indicated no significant effects of age or gender.

During production of the syllable /pi/, 2 young females, 3 older females, 3 young males, and 3 older males were categorized as exhibiting incomplete velopharyngeal closure. Data for /pi/ were not obtained for five young female speakers. During production of the syllable /si/, 0 young females, 3 older females, 3 young males, and 2 older males were categorized as exhibiting incomplete velopharyngeal closure. Data for /si/ were not obtained for 1 young male and 5 young females. Results of the chi squares indicated no significant effects of age or gender on the number of subjects categorized as having either complete or incomplete velopharyngeal closure.

Means and standard deviations of the ratios of nasal air volume as a function of age and gender during production of ''hamper'' are presented in Table II. Results of the ANOVAs indicated no significant effects of age or gender for the ratios associated with either the segmental and word boundary volumes or just the segmental volumes.

B. Declination effects

Figure 3 illustrates means and standard deviations of P_0 associated with /p/ in ''hamper'' as a function of age, gender, and word position in the utterance. Mean P_0 values associated with the second, third, and fourth repetitions of ''hamper'' for all speakers were 5.9 (s.d.=1.6), 5.6 (s.d.=1.6), and 5.6 (s.d.=1.6) cm H_2O. Results of the repeated measures ANOVA revealed a significant main effect of word

TABLE II. Means and standard deviations (s.d.) of nasal air volume ratios associated with ''hamper'' as a function of age and gender.

| | Young speakers | | | | Older speakers | | | |
| | Male | | Female | | Male | | Female | |
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Ratios								
Segmental	0.24	0.14	0.29	0.12	0.22	0.08	0.25	0.08
Segmental and word boundary	0.50	0.21	0.50	0.15	0.53	0.18	0.48	0.18
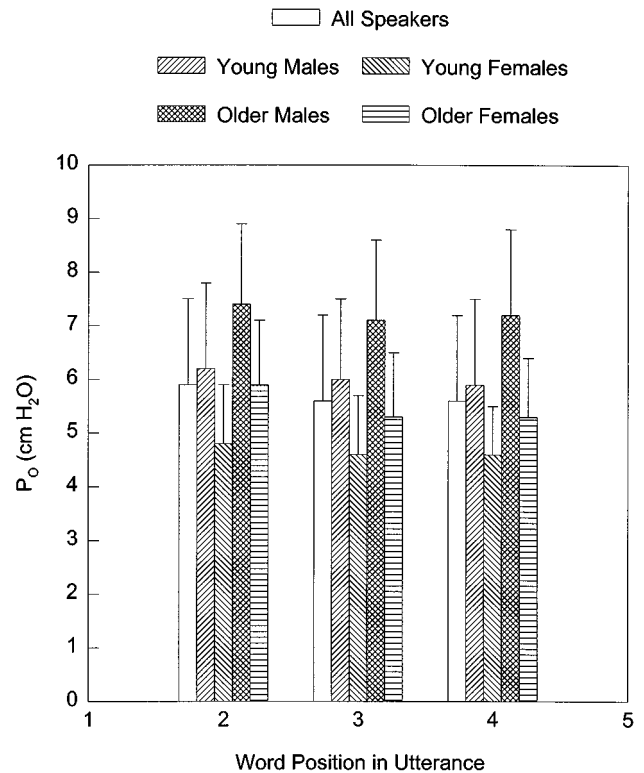
FIG. 3. Means and standard deviations of intraoral air pressure (P_0) of /p/ in ''hamper'' as a function of age, gender, and word position in the utterance.

position for all speakers ($F=10.4$, $df=118$, $p=0.0001$); main effects of age and gender were not significant. *Post hoc* analyses (Tukey HSD tests) indicated that the P_0 differences between the second and third repetitions and between the second and fourth repetitions were significant ($p<0.05$).

Figure 4 illustrates means and standard deviations of the ratios of nasal air volume associated with the segmental and word boundary effects of ''hamper'' as a function of age, gender, and word position in the utterance. Results of the repeated measures ANOVA revealed a significant gender by word position interaction ($F=4.0$, $df=2$, $p=0.02$). *Post hoc* analysis (Tukey HSD tests) of simple effects by gender revealed significant differences for female speakers between the second (mean=0.46, s.d.=0.22) and third (mean=0.51, s.d.=0.17) repetitions and between the second (mean=0.46, s.d.=0.22) and fourth (mean=0.53, s.d.=0.15) repetitions of ''hamper'' ($p<0.05$).

Figure 5 illustrates means and standard deviations of the ratios of nasal air volume associated with just the segmental effects of ''hamper'' as a function of age, gender, and word position in the utterance. Results of the repeated measures ANOVA revealed a significant gender by word position interaction ($F=3.3$, df=2, $p=0.04$). *Post hoc* analysis (Tukey HSD tests) of simple effects by gender revealed significant differences for male speakers between the second (mean=0.27, s.d.=0.16) and fourth (mean=0.23, s.d.=0.12) repetitions and between the third (mean=0.26, s.d.=0.13) and fourth (mean=0.23, s.d.=0.12) repetitions of ''hamper'' ($p<0.05$).
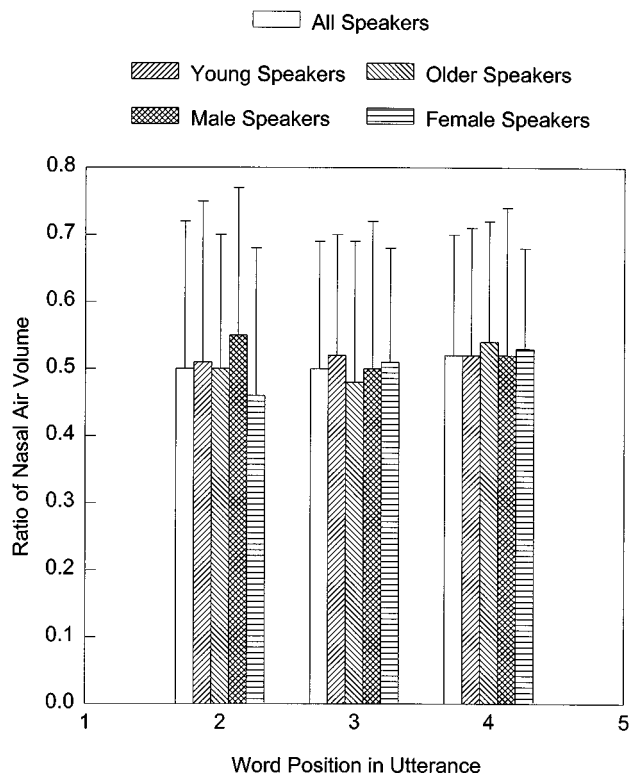
FIG. 4. Means and standard deviations of the ratios of nasal air volume associated with the segmental and word boundary effects of "hamper" as a function of age, gender, and word position in the utterance.

FIG. 5. Means and standard deviations of the ratios of nasal air volume associated with segmental effects of "hamper" as a function of age, gender, and word position in the utterance.

Figure 6 illustrates means and standard deviations of nasal air volumes associated with the /m/ segment in "hamper" as a function of age, gender, and word position in the utterance. Results of the repeated measures ANOVA revealed no significant age, gender, or word position effects.

III. DISCUSSION

The primary purpose of this study was to investigate velopharyngeal function in young and older adult speakers. The results indicated that, in general, the older speakers did not show any aerodynamic evidence of deterioration of velopharyngeal function. Older speakers exhibited similar levels of nasal airflow as young speakers during production of the oral consonant /p/ in a nasal-plosive sequence. Similarly, older speakers exhibited comparable ratios of anticipatory nasal air volume during "hamper." The hypothesis that older speakers would exhibit a velopharyngeal declination effect during a reiterant speech task was not supported by the nasal air volume data. When word position in the utterance was systematically evaluated, the older speakers performed similarly to the young speakers relative to anticipatory nasalization. Older speakers, however, exhibited greater levels of intraoral air pressure than young speakers.

The nasal airflow and volume findings from the word "hamper" provide additional evidence that the aerodynamic integrity of the velopharyngeal structures does not appear to deteriorate with age. The nasal airflow data from the syllables /pi/ and /si/ also support this statement. These findings are in agreement with Hoit et al. (1994) who also observed

essentially airtight velopharyngeal closure in older speakers. Analysis of the sequential measurements additionally suggest that older speakers are not susceptible to velopharyngeal declination within a breath group, at least for the word positions evaluated. This finding continues to be at odds, however, with previous acoustic data which showed that nasalance tended to increase as a function of speaking duration in older speakers (Hutchinson et al., 1978). One explanation that may help elucidate this apparent discrepancy involves declination of other supraglottic articulators. As Vayra and Fowler (1992) have shown, jaw opening may systematically decrease as a function of utterance length. Because nasalance is computed as a ratio of nasal to oral plus nasal acoustic energy, a decrease in oral cavity impedance from beginning to end of an utterance may result in progressive nasalance. This explanation, however, would also predict progressive nasalance for young speakers, which was not reported by Hutchinson et al. (1978).

The finding that older speakers exhibited greater intraoral air pressure levels during production of /p/ in "hamper" as compared to young speakers was not unexpected. This finding suggests differences in respiratory effort levels between young and older speakers. Indeed, Hoit and Hixon (1987) reported differences in speech breathing as a function of age in adult men. Older men (mean age=75 yr) were reported to initiate extemporaneous utterances at higher rib cage volumes than younger men (mean age=25 yr). The finding of higher intraoral air pressure levels in older speakers, therefore, is consistent with the known relationship between lung volume initiation levels
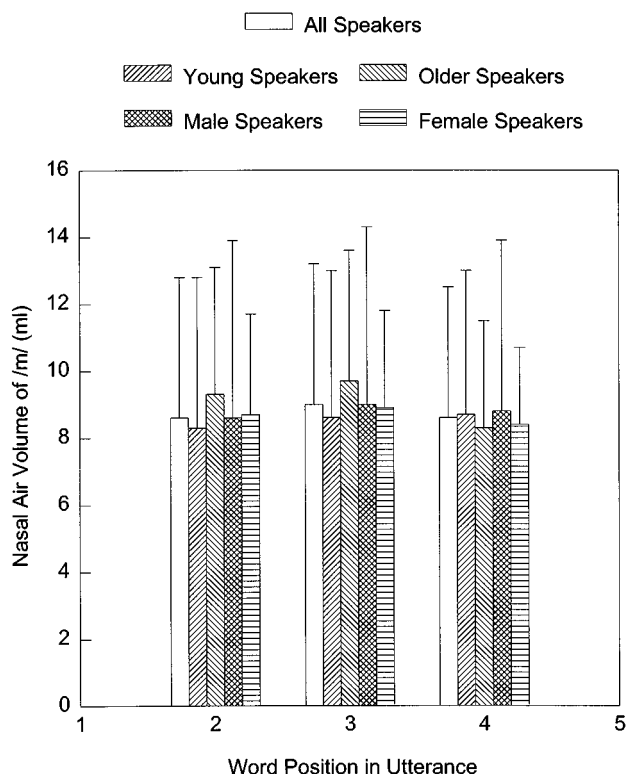
FIG. 6. Means and standards deviations of nasal air volume (ml) associated with /m/ of "hamper" as a function of age, gender, and word position in the utterance.

and relaxation pressures generated by elastic recoil. Although subglottal pressure and lung volume levels may be independently manipulated during speech production, older individuals may be less efficient at this task due to decreased compliance of the lungs (Kahane, 1990). It may also be argued, however, that age-related hearing loss in the elderly may have resulted in elevated sound-pressure levels and therefore increased subglottal air pressure levels. Although possible, it must be emphasized that no older speaker reported significant hearing loss and that all older speakers were judged to exhibit appropriate vocal loudness during perceptual screening procedures. In addition, the magnitude of the intraoral air pressure difference between the young and older speakers was approximately 1 cm H_2O. This is approximately the same magnitude of pressure difference that occurred between young males and females, suggesting lung volume initiation levels as a common mechanism. Men, for example, are known to exhibit greater relaxation pressures than women even at similar lung volume levels (Bode et al., 1976).

Additional results of the study indicated that (a) male speakers exhibited greater levels of intraoral air pressure than females, (b) all speakers exhibited declination of intraoral air pressure as a function of word position in the utterance, and (c) male and female speakers exhibited different patterns of nasal airflow as a function of word position in the utterance. The finding that male speakers, regardless of age, exhibited greater intraoral air pressure levels than female speakers was not unexpected. Zajac and Mayo (1996) reported this effect for the young speakers studied in the present investigation.

As suggested by Zajac and Mayo (1996), this finding may be explained by differences in respiratory function (described above) and/or preferred sound-pressure levels between men and women.[2] Also, declination of intraoral air pressure within a breath group was not unexpected. Indeed, declination of subglottal pressure and fundamental frequency of voice within an utterance are well-known phenomena (e.g., Honda, 1995). As indicated by Krakow (1993), these declination effects may be explained by a decrease in energy which occurs later in an utterance. It must be noted that the relatively small magnitude of pressure declination in the present study may have occurred, in part, due to the measurement procedures employed. Because the initial and final productions of the words and syllables were excluded from analysis to avoid nasal airflow artifact, this may have reduced the overall pressure declination effect.

The findings of gender specific nasal air volume characteristics as a function of word position in the utterance were unexpected. Krakow et al. (1995) suggested that velar declination was a general characteristic of speech production, similar to declination effects observed for other articulators. However, the present findings do not entirely support this notion. Indeed, only female speakers appeared to exhibit a velopharyngeal declination effect as reflected by an overall increase in nasalization during production of "hamper." This may have occurred as a result of biomechanical factors involving movement and inertia. McKerns and Bzoch (1970), for example, noted differences in the patterns of velopharyngeal closure between men and women. Women were reported to achieve a greater degree of contact of the velum against the posterior pharyngeal wall. This pattern was described as a distinctive two-stage movement of the velum (upward and backward) as compared to a more unidirectional movement exhibited by men. Consistent with this finding, Kuehn (1976) reported longer velar movement durations for a female as compared to a male speaker. It has also been reported that women may engage in earlier anticipatory nasalization than men (Thompson and Hixon, 1979). Zajac and Mayo (1996) further reported that women required longer durations to achieve peak oral pressure in the context of a nasal-plosive sequence. Viewed collectively, these studies suggest that biomechanical factors may facilitate velar declination in women, at least at the phrase level of production. The finding of decreased anticipatory nasal airflow as a function of word position in the utterance by the male speakers, however, is more difficult to explain. This result suggests that velopharyngeal closure actually increased during the vowel segments. If men typically achieve less velopharyngeal closure than women (Zajac and Mayo, 1996), then men may compensate during a reiterant speech task by increasing closure. Regardless of the reason, this finding does not support the notion of velar declination as a general phenomenon of speech production. Finally, it must be acknowledged that (a) the gender specific findings were relatively small in magnitude, especially for the male speakers, and (b) the lack of aerodynamic findings suggesting velopharyngeal declination as a general phenomenon does not rule out the presence of a velar kinematic effect. Such an effect, if present, may not have generated aerodynamic consequences.

Additional research, therefore, is required to clarify the relationships among kinematic, aerodynamic, and acoustic consequences of velar function.

IV. CONCLUSION

The results of the present study provide additional information relative to velopharyngeal function across a relatively wide age range of adult speakers. Consistent with previous aerodynamic studies, velopharyngeal function appeared not to deteriorate with age. The finding of a gender specific velopharyngeal declination effect suggests the involvement of biomechanical factors at the phrase level of production.

ACKNOWLEDGMENTS

[1]An unfiltered nasal airflow signal that was obtained with a partitioned circumferentially vented mask is illustrated in Fig. 2. This was done to emphasize the voicing component and its relationship to the acoustic signal.

[2]The author acknowledges that monitoring sound-pressure levels of the speakers would have facilitated interpretation of various aspects of the data. However, it must be emphasized that (a) the focus of the study was on the aerodynamic aspects of oral and nasal consonants, (b) the pressure-flow method, because of the use of oral and nasal catheters, may interfere with the recording of sound-pressure levels, especially for nasal sounds, and (c) at the time of data collection, the PERCI-PSCOPE system was limited relative to sampling rates for acoustic data.

Bode, F. R., Dosman, J., Martin, R. R., Ghezzo, H., and Macklem, P. T. (**1976**). ''Age and sex differences in lung elasticity and in closing capacity in nonsmokers,'' J. Appl. Physiol. **41**, 129–135.

Hoit, J. D., and Hixon, T. J. (**1987**). ''Age and speech breathing,'' J. Speech Hear. Res. **30**, 351–366.

Hoit, J. D., Watson, P. J., Hixon, K. E., McMahon, P., and Johnson, C. L. (**1994**). ''Age and velopharyngeal function during speech production,'' J. Speech Hear. Res. **37**, 295–302.

Honda, K. (**1995**). ''Laryngeal and extralaryngeal mechanisms of f_0 control,'' in *Producing speech: Contemporary issues for Katherine Safford Harris*, edited by F. Bell-Berti and L. J. Raphael (American Institute of Physics, Woodbury, New York).

Hutchinson, J. M., Robinson, K. L., and Nerbonne, M. A. (**1978**). ''Patterns of nasalance in a sample of normal gerontologic subjects,'' J. Commun. Disord. **11**, 469–481.

Kahane, J. (**1990**). ''Age-related changes in the peripheral speech mechanism: Structural and physiological changes,'' ASHA Reports **19**, 75-87.

Krakow, R. A. (**1993**). ''Nonsegmental influences on velum movement patterns: Syllables, sentences, stress, and speaking rate,'' in *Phonetics and Phonology: Nasals, Nasalization, and the Velum*, edited by M. Huffman and R. A. Krakow (Academic, San Diego).

Krakow, R. A., Bell-Berti, F., and Wang, Q. E. (**1995**). ''Supralaryngeal declination: Evidence from the Velum,'' in *Producing Speech: Contemporary Issues for Katherine Safford Harris*, edited by F. Bell-Berti and L. J. Raphael (American Institute of Physics, New York).

Kuehn, D. P. (**1976**). ''A cineradiographic investigation of velar movement variables in two normals,'' Cleft Palate J. **13**, 88–103.

McKerns, D., and Bzoch, K. R. (**1970**). ''Variations in velopharyngeal valving: The factor of sex,'' Cleft Palate J. **7**, 652–662.

Seaver, E. J., Dalston, R. M., Leeper, H. A., and Adams, L. E. (**1991**). ''A study of nasometric values for normal nasal resonance,'' J. Speech Hear. Res. **34**, 715–721.

Thompson, A. E., and Hixon, T. J. (**1979**). ''Nasal air flow during normal speech production,'' Cleft Palate J. **16**, 412–420.

Vatikiotis-Bateson, E., and Fowler, C. (**1988**). ''Kinematic analysis of articulatory declination,'' J. Acoust. Soc. Am. Suppl. 1 **84**, S128 (A).

Vayra, M., and Fowler, C. (**1992**). ''Declination of supralaryngeal gestures in spoken Italian,'' Phonetica **49**, 48–60.

Warren, D. W., Dalston, R. M., Trier, W. C., and Holder, M. B. (**1985**). ''A pressure-flow technique for quantifying temporal patterns of palatopharyngeal closure,'' Cleft Palate J. **22**, 11–19.

Zajac, D. J., and Mayo, R. (**1996**). ''Aerodynamic and temporal aspects of velopharyngeal function in normal speakers,'' J. Speech Hear. Res. **39**, 1199–1207.

Expression of emotional–motivational connotations with a one-word utterance

Lea Leinonen and Tapio Hiltunen
Neural Networks Research Centre, Helsinki University of Technology, Rakentajanaukio 2 C, 02150 Espoo, Finland

Ilkka Linnankoski and Maija-Liisa Laakso
Department of Biomedicine, University of Helsinki, Siltavuorenpenger 20 J, 00014 University of Helsinki, Finland

The expression of ''naming,'' ''commanding,'' ''angry,'' ''frightened,'' ''pleading,'' ''astonished,'' ''satisfied,'' ''admiring,'' ''scornful,'' and ''sad'' was studied with the word [saara] spoken by 12 subjects. Using the same connotations, the 120 utterances were categorized by 73 listeners. Most samples were agreed on by 50%–99% of the judges. Most samples of ''astonished,'' ''angry,'' ''frightened,'' and ''commanding'' were judged as intended, while ''pleading'' was often confused with ''sad,'' and ''content'' with ''admiring.'' Acoustic differences between the categories were examined for $F0$, duration, and sound pressure; spectral features of [aa] were visualized with the self-organizing map of Kohonen. Most intraspeaker variation of mean $F0$, $F0$ range, sound pressure, and duration took place during the [aa] segment. Peak sound pressure, mean $F0$, and spectral energy distribution of [aa] differentiated among ''commanding,'' ''angry,'' ''frightened,'' ''naming,'' and ''sad.'' Specific intonations of the [aa] segment were encountered for ''astonished,'' ''scornful,'' and ''pleading.'' The best-conveyed ''admiring'' samples were distinguished from ''content'' by spectral cues for a breathy voice quality. © *1997 Acoustical Society of America.* [S0001-4966(97)02609-X]

PACS numbers: 43.70.Fq, 43.70.Gr, 43.71.Bp [AL]

INTRODUCTION

Speaker affect modifies phonation, articulation, and facial expression, which thus may all contribute to the acoustic signaling of the speaker's emotional–motivational state. From an evolutionary and ontogenetic perspective, acoustic signaling of emotions primarily relies on phonation but becomes modified by the concomitant facial display: jaw opening and lip retraction/protrusion. Humans share the neurobiological basis of vocal communication with other primate and mammalian species. The behavioral attributes of these homologies are, however, difficult to determine because of the diversity of body shapes and vocalizations. Our interest in shared features of primate vocalizations was evoked by the observation that naïve human listeners were able to identify emotional messages in the vocalizations of macaque monkeys (*Macaca arctoides*; Leinonen *et al.*, 1991), and that the ability to interpret monkey vocalizations developed during childhood (Linnankoski *et al.*, 1994) in a way similar to the recognition of human emotions (McCluskey *et al.*, 1975; Fenster *et al.*, 1977; van Bezooyen, 1984). The results suggested that monkeys and man share acoustic cues for fear, aggression, dominance, submission/plea, satisfaction, and emotional neutrality. To test the hypothesis, we needed human samples of about the same duration as the monkey vocalizations. Affect sounds, functionally resembling the monkey sounds, have a very restricted role in adult communication. We therefore decided to explore whether a single word would allow perceptual and acoustic differentiation among the same connotations that were identified in the

monkey vocalizations. To facilitate the analysis, and the speakers' task, we used a carrier word which would be used naturally for emotive signaling. A person's name was considered suitable because speakers often express their emotional–motivational state when uttering the listener's name. The name [saara] (the Finnish equivalent to Sarah) was selected because the long vowel made it suitable for comparison with monkey vocalizations.

The word [saara] was spoken simulating 10 emotional–motivational connotations. Six of the connotations, ''angry,'' ''frightened,'' ''commanding,'' ''pleading,'' ''content,'' and emotionally neutral ''naming'' corresponded to those used in the monkey experiments. We also included some connotations used for human emotions ''astonished'' (positive surprise, not just an orienting reflex with vegetative responses and arousal), ''scornful,'' ''admiring,'' and ''sad.'' Except for ''pleading,'' ''admiring,'' and ''commanding'' the categories have been frequently used in studies on vocal emotive expression in man (Scherer, 1986, 1995; Murray and Arnott, 1993).

Human emotions have usually been studied with running speech (often single sentences) which allows differentiation among at least ten connotations (Davitz and Davitz, 1959; Davitz, 1964; van Bezooyen, 1984). Whether a single word could signal as many connotations has not been studied, but the communication of six emotions by an isolated vowel has been suggested by Kaiser (1962). Expressing emotions with a single word instead of with running speech may result in exaggeration and in the reduction of inter- and intraspeaker

variation that may facilitate the identification of meaningful signal dimensions.

In emotive speech, significant variation has been observed for fundamental frequency (level, variation, and contour), amplitude, duration, tempo, and voice spectrum (reviewed recently by Murray and Arnott, 1993; and Scherer, 1995), and it has been suggested that the level of fundamental frequency, intonation contour, and voice quality vary independently of each other (Ladd *et al.*, 1985). The means of emotive expression are also used for semantic differentiations among individual words and sentences, and the separation of these two sources of variation in the speech signal is not easy. Most of the language-related variation is abolished by the use of a single word. The present study was carried out on native speakers of Finnish in which the main stress is always on the first syllable of a word, sentence intonation is generally falling (also in questions), and the difference between long ([aa] in the first syllable of [saara]) and short ([a] in the second syllable) phonemes is used to distinguish different words (Karlsson, 1987).

I. METHOD

A. Speech samples

Eight men and eight women, 23–47 years of age (students and teachers of medicine, economics, logopedics, and engineering), were asked to produce [saara] with 10 different connotations defined by 10 frame stories. The speakers were asked to read the stories and identify themselves with the person who addressed Saara for different reasons. The sex of the speaker was not defined. For each category, several samples of [saara] were recorded. Three samples for each speaker and each category were edited onto a listening tape. To avoid inclusion of too many "poor" samples, which would not allow differentiation among the categories, a preselection was carried out with six listeners. They were asked to indicate the best sample (out of three) for each category, and whether the selected best sample was "very bad" with respect to the category. The samples of 4 speakers were more than 11 times out of 60 indicated as "very bad" and these speakers were discarded from the study. The perceptual categorization test and the acoustic analyses were carried out with the best samples (one for each category) of seven women and five men.

The 10 categories were defined for the speakers, and later for the listeners, as "stating, informing, naming" (in Finnish "toteava, tiedottava, nimeävä"; hereafter referred to as "naming"); "astonished, surprised" ("hämmästynyt, yllättynyt"; "astonished"); "pleading, submissive" ("anova, alistuva"; "pleading"); "content, pleased, happy" ("tyytyväinen, mielissään, iloinen"; "content"); "admiring, adoring" ("ihaileva, palvova"; "admiring"); "sad, depressed" ("surullinen, masentunut"; "sad"); "commanding, threatening, domineering" ("käskevä, uhkaava, määräilevä"; "commanding"); "contemptuous, scornful" ("ylenkatsova, halveksiva"; "scornful"); "frightened, terrified" ("pelästynyt, kauhuissaan"; "frightened"); and "angry, cross, furious" ("vihainen, kiukkuinen, raivoissaan"; "angry"). The frame stories for the different categories are briefly sum-

marized as follows. The "naming" [saara] was spoken by the chairman of a meeting who asked for comments and named Saara who had raised her hand. The "astonished" [saara] was spoken by a music teacher who met her/his pupil unexpectedly during a summer vacation abroad. The "pleading" [saara] was spoken by Saara's young colleague who wanted the reluctant Saara to take the evening duty so that she/he could meet her/his young love. The "content" [saara] was spoken by Saara's sister/brother who had felt tired and cold after walking for hours in the rain and slush and was now sitting comfortably by the fireside and having tea served by Saara. The "admiring" [saara] was spoken by Saara's cousin when she/he saw young Saara in her new wonderful ball dress. The "sad" [saara] was spoken by Saara's friend who had not passed the final examination in spite of all the hard work. The "commanding" [saara] was spoken by a strict and authoritarian camp leader who wanted Saara to stop talking when he/she was giving orders. The "scornful" [saara] was spoken by a person with whom Saara tried to make friends after deceiving her/him. The "frightened" [saara] was spoken by her friend when they met a dangerous dog which started to charge toward them on a path in the woods. The "angry" [saara] was spoken by her roommate when seeing her/his carefully sorted stamps pushed to the corner of the table by a nonchalant Saara.

B. Recording and signal sampling

The speech samples were recorded with a microphone at a distance of 30 cm from the mouth (AKG C568EB and Sony TCD-D3 DAT recorder) in an anechoic room. The original sound pressure levels at the microphone were measured from the recorded samples with the aid of a calibration signal. The analog signals were digitized in an Iris Indigo workstation at a 22.05-kHz sampling frequency and 16-bit quantization. The peak loudness of the neutral utterances varied with speaker up to 20 dB (A) (Brüel & Kjaer 2232) and the intraspeaker variation in emotive speech was up to 45 dB (A). The amplitudes were adjusted before digitization so that the peak loudness varied randomly within 5 dB (A).

C. Auditory rating test

The digitized samples of the speakers were randomly (the male and female samples were not mixed) edited onto a listening tape: each sample was repeated four times, followed by a 5–6 s pause before the next one. The tape included random repetitions of three female and seven male samples. The rating test was carried out in an ordinary lecture room and the judges were at a distance of 2–4 m from two loudspeakers. Six groups were tested, altogether 73 judges, 46 women and 27 men, who were students of medicine, psychology, and engineering, and personnel of medical and technical research laboratories. The judges were asked to choose 1 of the 10 categories for each sample. The test with the female samples lasted 19 min and the test with the male samples 15 min.

In the listening tape, the peak amplitudes of the samples varied within 5 dB (A). This assured that even the originally weakest utterances were audible to everybody and that the

judges were not startled by extremely loud utterances. Variation in the speaker's effort, measured with sound pressure at the microphone, is associated with changes in the glottal waveform (Holmberg *et al.*, 1988; Cummings and Clements, 1995). The audible cues of the signal waveform enable the listener to judge correctly the speaker's effort to sound pressure irrespective of the playback intensity of a recorded signal (Brandt, 1972). Sound pressure at the microphone may thus be regarded as an index of effort-related variation in the distribution of spectral energy. In the present study it was assumed that the emotional speech mode is also signaled by acoustic cues which an adult listener can correctly judge irrespective of the playback intensity, or the distance from the (loud) speaker.

D. Acoustic analyses

The following measurements were made from the recorded and digitized signals: duration of [s], [aa], and [ra], mean $F0$ during the [aa] and [a] segments, $F0$ range during the [aa] segment (difference between the maximum and minimum), A-filtered peak sound pressure level of [aa], and peak level differences between [s] and [aa], and [aa] and [ra]. A-filtered sound pressure curves of [saara] samples, $F0$ contours during [aa], and spectral models for [aa] obtained with the self-organizing map, a neural network algorithm (Kohonen, 1995), were evaluated visually.

The durations were manually measured from the digitized signals, accuracy was of the order of 10 ms. The A-filtered peak sound pressure level during [aa] was measured from the playback of the DAT records using a 125-ms time constant (Brüel & Kjaer 2232); the recording level was determined with the aid of a recorded 1-kHz calibration tone. A-filtered sound pressure curves of [saara] and peak pressure differences between [s] and [aa], and [aa] and [ra] were obtained from A-weighted Hamming-windowed 23-ms Fourier spectra, computed at 10-ms intervals. $F0$ was determined using a combination of autocorrelation and cepstral analysis (Rihkanen *et al.*, 1994). The mean $F0$ was calculated for the [aa] and [a] segments of [saara]. $F0$ contours of the [aa] segments were plotted for visual evaluation.

The self-organizing map (Kohonen *et al.*, 1992; Kohonen, 1995; applied to clustering of speech spectra, e.g., by Leinonen *et al.*, 1993 and Rihkanen *et al.*, 1994) was used for the visualization of distinguishing spectral patterns. The self-organizing map creates, unsupervised from examples, a topologically ordered representation of the sample space by taking the probability distribution into account. A two-dimensional map consists of a number of sample models. During the training of the map, these models come to represent the recurrent features in input patterns. The location of a sample on a trained map is determined by the best-matching model (according to Euclidean distance); as nearby models are similar, similar samples are usually projected near to each other on the map.

In the present study, spectral samples were obtained from Hamming-windowed 23-ms Fourier spectra, computed at 10-ms intervals with 10-dB attenuation at the low end. The spectra were smoothed with 15 truncated sinc-function filters with the following center frequencies and widths of the cen-

ter window (in parentheses): 0.1 kHz, 0.15 kHz, 0.2 kHz, 0.3 kHz, 0.35 kHz, 0.45 kHz (0.1 kHz), 1 kHz (0.5 kHz), 2 kHz, 3 kHz (1 kHz), 3.5 kHz, 4 kHz, 4.5 kHz, and 5 kHz (0.5 kHz), 6 kHz, 7 kHz (1 kHz). The filter outputs constituted a 15-component spectral feature vector fed to a map. Maps of 18×20 locations, each location with a 15-component model, were trained separately with all male and all female samples. The best-matching model vectors for the samples on which over 50% of the listeners agreed were used to visualize the intercategory differences in spectral patterns.

II. RESULTS

A. Judges' performance

The reliability of the listener group was studied with 10 randomly selected repeated samples. A statistically significant difference ($p<0.01$, chi-square test) was found for one sample: it was regarded on both trials as "angry" and "commanding" but the distribution of the choices was somewhat different. The variation of the judges performance was thus probably due to fluctuation of the decision border with the previously heard samples. Both female and male judges identified "correctly" (as intended) 50% of the samples, 42% of the male and 56% of the female samples (in these respects there were no significant differences between the female and male judges; Mann–Whitney U test).

B. Intercategory differentiation by the judges

None of the speech samples were categorized randomly. As shown in Table I, the connotations "commanding," "astonished," "angry," and "frightened" were conveyed best. "Pleading" was often regarded as "sad," "sad" as "pleading," "content" as "admiring," "naming" and "angry" as "commanding," and "scornful" as "naming." For the male speakers it was especially difficult to convey the messages of "pleading," "content," "admiring," and "sad": 65% of the male samples but only 43% of the female samples were not categorized as intended (the majority of judges did not choose the intended category). The emotionally rather neutral "naming" category was chosen for 13% of the male and for 5% of the female samples with other intents.

All categories, including those with low average recognition (Table I), included samples which were agreed on by over 70% of the judges. The judges' agreement was especially high, 85%–99%, for some "angry," "commanding," "astonished," and "scornful" utterances of both female and male speakers, and for some "frightened" utterances of female speakers. The judges' agreement was higher for the female than male samples. For instance, 76% of the female but only 54% of the male samples were agreed on by over 50% of the judges.

None of the samples were agreed on by all listeners. Cooccurrences of listener choices in pairs of categories were measured with Spearman's correlation coefficients; Table II. (Nonparametric tests were used in the present study because populations were not normally distributed and had different variances.) The table shows that differentiation between "commanding" and "angry," between "content" and "admiring," and between "pleading" and "sad" was especially

TABLE I. Confusion matrix for 10 categories simulated by 12 speakers and evaluated by 73 judges. The values in each row show the distribution of listeners' choices (in percentages) for the 12 samples in each simulated category.

Simulated category	Chosen category									
	nam	comm	angry	scorn	cont	admir	plea	sad	fright	ast
naming	**48**	29	1	5	10	1	3	0	0	3
commanding	6	**75**	14	4	0	0	0	0	0	1
angry	0	22	**70**	3	1	0	1	0	2	1
scornful	18	6	2	**52**	4	2	5	8	1	2
content	9	2	0	5	**26**	22	10	12	1	13
admiring	10	2	0	8	9	**34**	16	10	3	8
pleading	11	3	1	3	1	2	**37**	33	2	7
sad	19	3	0	10	4	4	23	**36**	0	1
frightened	3	4	1	1	3	3	9	3	**64**	9
astonished	5	3	1	4	11	1	2	1	0	**72**

difficult. ''Scornful'' was confused only with the emotionally neutral ''naming'' suggesting that the acoustic cues of ''scornful'' were quite specific. The table also shows that ''commanding'' was better differentiated from ''content,'' ''admiring,'' ''pleading,'' and ''sad'' in the male than in the female productions.

C. Relationships among [s], [aa], and [ra]

Significant duration differences were encountered for [aa] in 25, for [s] in 19, and for [ra] in 10 of 45 pairwise intercategory comparisons ($p < 0.02$, Friedman test with Wilcoxon test). Thus the duration of [ra] varied less with different connotations than the duration of [aa] ($p < 0.005$, chi-square test) or [s] ($p < 0.05$). Durations of [s], [aa], and [ra] varied usually in parallel; there were thus significant positive correlations among the durations of all segments ($p < 0.01$,

Spearman's r). For ''frightened,'' however, the duration of [ra] relative to [saa] was longer than for any other connotation ($p < 0.001$, Friedman test with Wilcoxon test).

The difference between the peak pressures of [s] and [aa] tended to increase with increasing pressure of [aa] ($p < 0.01$, Spearman's r). For ''scornful,'' the difference between [s] and [aa] relative to the peak pressure of [aa] was smaller than for any other simulated category ($p < 0.03$, except for admiring $p < 0.05$). The peak sound pressure was usually lower at the second than first syllable; exceptions were observed in 8% of the samples. The peak pressure difference between [aa] and [ra] did not correlate with the peak pressure of [aa] (Spearman's r). The pressure drop was most pronounced in the ''scornful'' samples and it was due to the fall of pressure during the first syllable; Fig. 1.

All [aa] samples were voiced, at least initially, whereas

TABLE II. Spearman's rank correlation coefficients between the shares of scores in 10 categories selected by 73 listeners for the samples of 7 female and 5 male speakers. The statistically significant ($p < 0.01$) positive correlations are in bold, the negative in italic.

Female speakers

	nam	comm	scorn	cont	admir	plea	sad	fright	angry	ast
naming	1.00									
commanding	0.29	1.00								
scornful	**0.31**	0.19	1.00							
content	**0.42**	−0.11	0.18	1.00						
admiring	0.14	−0.28	0.19	**0.52**	1.00					
pleading	0.11	*−0.55*	0.05	0.18	**0.45**	1.00				
sad	0.13	*−0.39*	0.29	0.04	**0.37**	**0.73**	1.00			
frightened	−0.01	0.19	0.08	−0.07	0.10	0.25	0.20	1.00		
angry	−0.11	**0.66**	0.18	−0.28	−0.21	*−0.35*	−0.13	0.27	1.00	
astonished	−0.06	−0.12	0.25	**0.30**	0.24	0.23	0.22	0.21	0.01	1.00

Male speakers

	nam	comm	scorn	cont	admir	plea	sad	fright	angry	ast
naming	1.00									
commanding	−0.02	1.00								
scornful	0.01	0.28	1.00							
content	**0.33**	*−0.31*	0.29	1.00						
admiring	0.08	*−0.57*	−0.03	**0.48**	1.00					
pleading	0.14	*−0.66*	−0.20	0.11	**0.62**	1.00				
sad	0.23	*−0.63*	−0.13	0.06	**0.57**	**0.83**	1.00			
frightened	0.15	0.27	0.09	0.14	0.29	**0.40**	**0.33**	1.00		
angry	−0.16	**0.67**	0.12	*−0.35*	−0.30	*−0.31*	−0.19	**0.31**	1.00	
astonished	0.06	−0.06	0.07	**0.41**	0.18	0.06	−0.13	**0.53**	−0.14	1.00
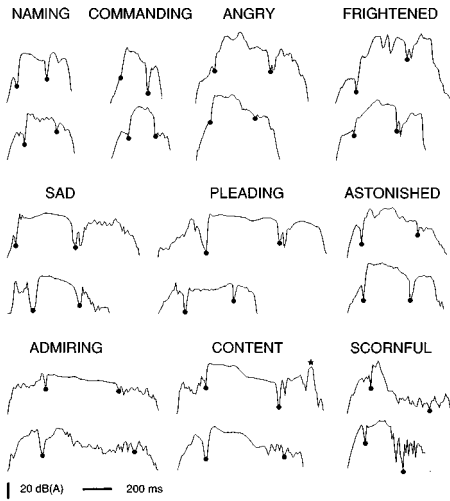
FIG. 1. Sound pressure contours, computed at 10-ms intervals from A-filtered Fourier spectra, for the two best-recognized female or male samples in each category (agreed on by 50%–99% of 73 listeners). The black dots on the curves mark the transition between [s], [aa], and [ra]. The black star at the end of one ''content'' curve indicates an [a] segment preceded by a transient cessation of voicing which was regarded as a laugh.

in 17% of the utterances the second syllable was unvoiced. Unvoicing was most frequent (8 of 12) in the simulations of ''scornful,'' in which the voicing ended already during the [aa] segment, but it was also encountered in ''admiring,'' ''content,'' ''pleading,'' ''sad,'' and ''naming.'' Significant differences among the simulated categories were observed for the mean $F0$ of [aa] in 60% and for the mean $F0$ of [a] in 33% of the comparisons ($p<0.02$, Wilcoxon test). The categories differing with respect to [a] also differed with respect to [aa], suggesting that the mean $F0$ of [aa] and [a] varied in parallel. The ''frightened'' samples with the highest mean $F0$ of [aa] also had the highest mean $F0$ during [a] ($p<0.02$ for all comparisons with the other categories). In 10 of the 120 samples the mean $F0$ of [a] was higher than the peak $F0$ of [aa]; this was most frequent in samples with ''naming'' intents (in 7 of 12). However, the best-conveyed ''naming'' intent and most samples of the other intents

judged as ''naming'' by the listeners had a falling word intonation.

D. Variation of sound pressure, mean $F0$, $F0$ range, and duration of [aa]

As a group, the female and male speakers were both rather homogeneous with respect to the acoustic parameters of [saara]; there were significant interindividual differences ($p<0.03$, Kruskall–Wallis test with Dunn test) only for duration. The women and men used similar sound pressures but the female utterances were of longer duration, higher $F0$, and greater $F0$ variability, ($p<0.01$, Mann–Whitney U-test). Figure 2 shows variation of acoustic parameters both with the categories simulated by the speakers and with the categories chosen by the listeners. For the simulated categories, each including samples of the same speakers, group means are given for combined female and male data. In the categories chosen by the listeners the number of samples and the share of female and male data varied and therefore the means of $F0$ and duration, with significant intergender differences, are shown separately for female and male speakers. Overwiew of the parameters in Fig. 2 demonstrates the covariation of the durations of [s], [aa], and [ra]. The figure also suggests that the sound pressure and mean $F0$ of [aa] tended to increase simultaneously. Correlation analysis both of female and male samples showed that mean $F0$ and $F0$ range increased with peak sound pressure and the relationships were independent of duration ($p<0.01$, Spearman's r; first- and second-order partial correlations did not differ from simple correlation).

''Angry,'' ''frightened,'' and ''commanding'' intents had higher and ''sad'' intents lower sound pressure than ''naming,'' Fig. 2. ''Angry'' was expressed with a higher sound pressure than any other connotation ($p<0.002$, Friedman test with Wilcoxon test). In comparison to ''naming,'' the speakers increased $F0$ for ''frightened,'' ''angry,'' and ''astonished,'' most for ''frightened'' which in this respect differed significantly from all other connotations ($p<0.01$). The range and rate ($F0$ range divided by duration; data not shown) of $F0$ change during [aa] were highest for
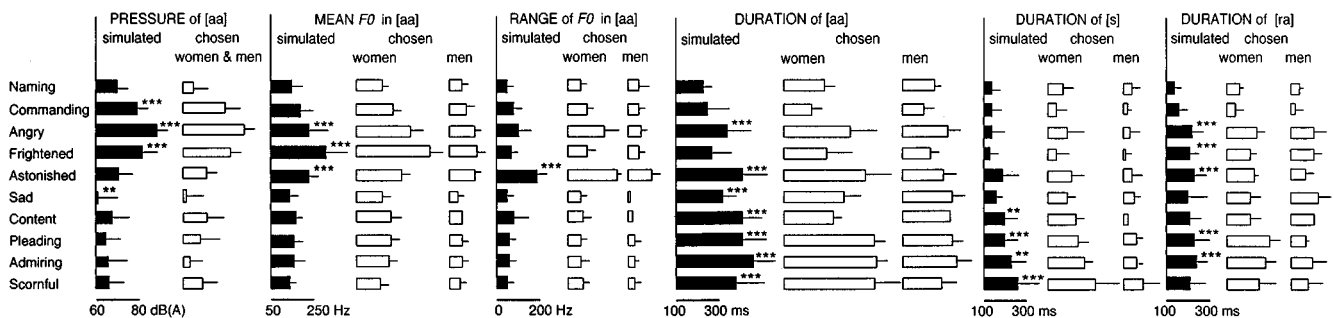
FIG. 2. Parameter means (with SD) for categories simulated by 12 speakers (black bars), and for categories chosen by 73 listeners (white bars). For the chosen categories, mean $F0$, $F0$ range, and sample duration, with significant intergender differences, are shown separately for women and men because the number speakers varies with category (''naming'' 8/9 (female/male), ''commanding'' 9/7, ''angry'' 7/4, ''frightened'' 7/4, ''astonished'' 7/7, ''sad'' 9/6, ''content'' 5/1, ''pleading'' 8/5, ''admiring'' 5/3, ''scornful'' 5/4). Parameter scales are given at the bottom of the figure (the scale for simulated and chosen categories is the same); sound pressure refers to that measured at the microphone. The asterisks indicate statistically significant differences between ''naming'' and the other categories ($p<0.02$ two asterisks, $p<0.002$ three asterisks, Friedman test with Wilcoxon test).

LOUD AND HIGH   SOFT AND LOW
pressure
highest   **lowest**

ang fri com ast   nam con sco adm plea sad

mean _F_0
highest   **lowest**

fri ang ast com   con adm plea nam sad sco

_F_0 range
highest   lowest   highest   lowest

ast ang com fri   con adm plea sco nam sad

duration
highest   lowest   highest   lowest

ast ang fri com   adm con plea sco sad nam

FIG. 3. Differences in peak sound pressure, mean F_0, F_0 range, and duration of [aa] among the simulated categories. Significant differences ($p < 0.02$, Friedman test with Wilcoxon test), shown by connecting lines, are given separately for ''loud and high'' and ''soft and low'' categories. The ''loud and high'' categories were of higher peak sound pressure (with the exception of ''astonished'') and of higher mean F_0 (with the exception of ''commanding'') than the ''soft and low'' categories ($p<0.02$ for all pairwise comparisons, Friedman test with Wilcoxon test).

○ naming
● commanding
● angry
✴ frightened
❧ astonished
◇ sad, plea, con, adm, sco

FIG. 4. A Sammon plot of four-dimensional acoustic feature vectors for the [aa] segment in the utterances of [saara] labeled according to the chosen category. The feature vector components were: peak A filtered sound pressure, mean F_0, F_0 range, and duration. All components were on scale 0–1 which was obtained by setting the female and male group variances to one. The dashed line was drawn to emphasize the relatively good separation between ''soft and low'' and ''high and loud'' categories. For clarity, only one symbol is used to label the overlapping ''sad,'' ''pleading,'' ''content,'' ''admiring,'' and ''scornful'' samples.

''astonished,'' which in these respects differed from all other categories ($p<0.03$). In comparison to ''naming'' and ''commanding,'' the samples of the other intents were of longer duration ($p<0.02$, for ''frightened,'' $p<0.03$). That ''commanding'' is conveyed by relatively short utterances was also corroborated by the observation that the mean duration of the samples judged as ''commanding'' by the listeners was somewhat shorter than that of the samples judged as ''naming;'' Fig. 2.

The statistical significance of variation with simulated category is shown in Fig. 3 for four dimensions: peak A-filtered sound pressure, mean F_0, F_0 range, and duration. To compress the representation, the categories are divided in the figure, using the significant differences in sound pressure and mean F_0, into ''loud and high'' and ''soft and low.'' All ''loud and high'' categories differed from each other either by peak sound pressure or mean F_0, whereas the ''soft and low'' categories were differentiated mainly by differences in duration; Fig. 3. ''Content,'' ''admiring,'' and ''pleading'' were not differentiated by any of the four parameters. The means of the samples categorized according to the listeners' interpretation (Fig. 2) are in accordance with this result: There are no differences in sound pressure, mean F_0, F_0 range, or duration that could be used by the listener to differentiate among ''content,'' ''admiring,'' and ''pleading.''

Since none of the subjects succeeded in conveying all connotations, differences among the chosen categories could not be evaluated using intraspeaker comparisons. If group means are compared, much of the variation revealed by intraspeaker comparisons becomes obscured due to the interspeaker differences in speech ranges. This is demonstrated by the Sammon plot (Sammon, 1969) of four-dimensional samples (peak A-filtered sound pressure, mean F_0, F_0 range, and duration of the [aa] segment) labeled according to

the category chosen by the majority of judges; Fig. 4. The Sammon mapping projects samples from the four-dimensional space onto a plane with optimum preservation of intersample distances. Figure 4 suggests that the group means distinguish among ''commanding,'' ''frightened-or-angry,'' ''astonished,'' and ''naming-or-sad-or-pleading-or-content-or-admiring-or-scornful'' categori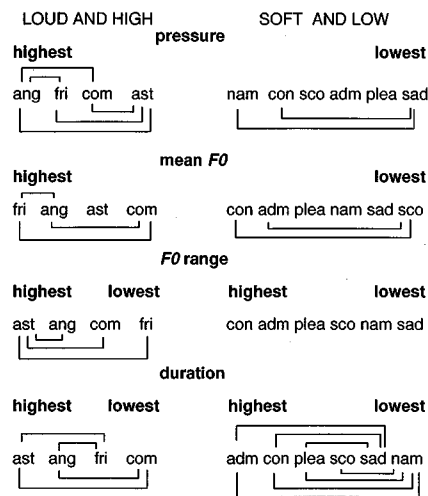es. ''Angry'' and ''frightened,'' for example, are not separated, although the intraspeaker comparisons in Fig. 3 showed that significantly higher F_0 was used for ''frightened'' than ''angry,'' and significantly higher pressure for ''angry'' than ''frightened.''

E. _F_0 contours of [aa]

There were 10 ''naming,'' 16 ''commanding,'' 10 ''angry,'' 8 ''frightened,'' 7 ''sad,'' 3 ''pleading,'' 4 ''admiring,'' 2 ''content,'' 12 ''astonished,'' and 8 ''scornful'' samples on which over 50% of the judges agreed. Visualization of the F_0 contours of the [aa] segments of these well-recognized samples showed intercategory differences which were only partly captured by the measures of mean F_0, F_0 range, and duration. These differences are illustrated in Fig. 5.

Rather stereotypic deviations from the habitual falling F_0 contour were observed for ''astonished,'' ''pleading,'' and ''scornful.'' All ''astonished'' samples had an initially high and rapidly falling F_0; Fig. 5. In ''pleading'' utterances, a slight rise of F_0 (and sound pressure; Fig. 1) was encountered toward the end of the [aa] segment. A rise of F_0 toward the end of the [aa] segment was observed only in three samples of the other categories and then it preceded an exceptional [ra] syllable with an even higher mean F_0. In ''scornful'' [aa], an initial voiced segment was followed by an extreme fall of F_0, resulting in cessation of phonation
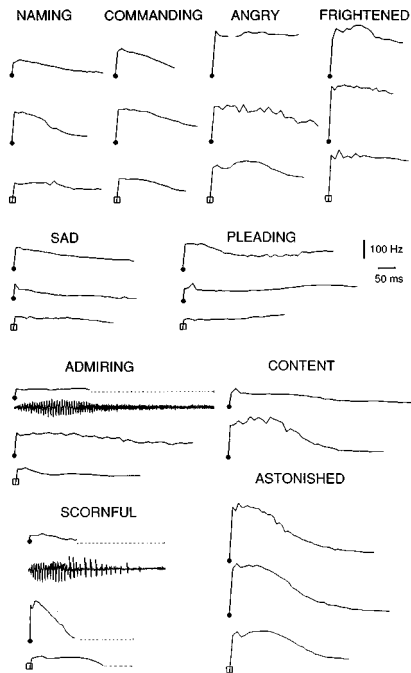
FIG. 5. F0 contours during the [aa] segment of [saara] in the samples on which most judges agreed. There are three samples for each category except for ''content'' (with no male sample agreed on by over 50% of judges). The traces of the female samples (marked with a black dot) begin at 100 Hz; those of men (marked with a rectangle) at 50 Hz to show level of F0 at the beginning of phonation. The audiosignals are shown for one ''admiring'' and one ''scornful'' sample in which the F0 curve breaks up.

which is demonstrated by the audio signal of one utterance in Fig. 5. This intonation pattern, with a concurrent extreme fall of sound pressure (Fig. 1) was not observed in any other category. The ''admiring'' connotations were characterized by breathy or whispery voice. The audio signal of one ''admiring'' sample in Fig. 5 demonstrates that this utterance ended in a whisper (without a preceding fall of F0 as in ''scornful'' samples).

F. Spectral patterns of [aa]

The spectral differentiation among the [aa] segments in the utterances agreed on by most listeners is visualized on self-organized spectral feature maps in Fig 6. The map distributions show grouping according to the connotation which suggests that some intercategory differences have been extracted during the training of the maps. Because there were more female than male samples and the female samples conveyed the connotations better, the spectral differentiation is clearer on the female than on the male map. On the female map, the ''angry'' and ''frightened'' samples and the first segments of the ''astonished'' samples are located with some overlap around the projection area of [s] samples. The ''commanding'' samples overlap with the ''angry'' and ''naming'' ones. The ''sad,'' ''pleading,'' and ''content'' samples and the second segment of the ''astonished'' samples overlap with the ''naming'' ones. Most of the ''scornful'' sample points are near the area for silence in the lower left corner of the map.
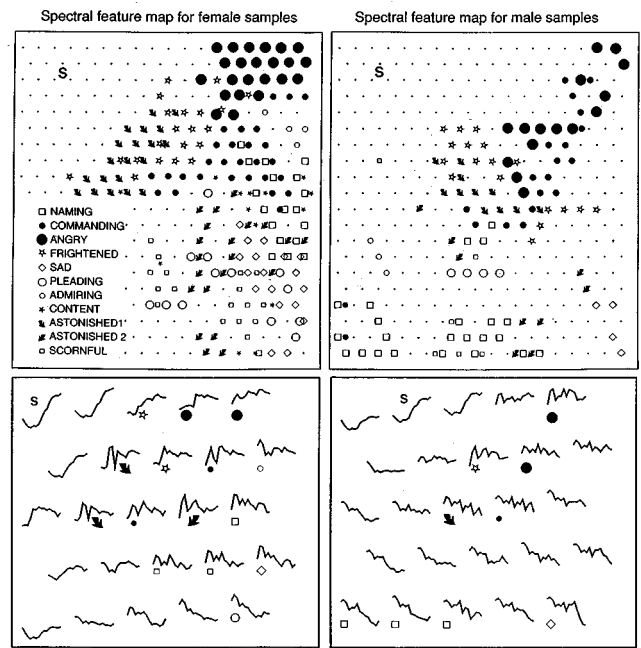
FIG. 6. Self-organized map projections of spectral samples of [aa] in utterances of [saara] categorized in agreement by over 50% of the judges (above). The small dots on the maps show the 360 locations, each corresponding to a model vector. The symbol [s] in the upper left corner of the maps shows the projection area of [s] segments. ''Astonished 1'' refers to the initial part of the [aa] segment with high F0, and ''astonished 2'' to the end part with much lower F0. The model vectors at every fifth location of the maps are shown below. The model vectors have 15 components corresponding to the 15 spectral filters used for the feature extraction. The seventh component in the middle of a model vector corresponds to the filter with the center frequency at 1 kHz. Thus the left half of a model vector represents energy distribution at frequencies below 1 kHz, and the right half above 1 kHz. The labels of the models have been taken from the map above.

The model vectors shown for every fifth of the 360 locations of the map (lower part of Fig. 6) differ with respect to the distribution of energy between the low and high end of the spectrum. The ''commanding'' models have relatively more energy at the main formants around 1–4 kHz than the ''naming'' models, and the shift of energy is even more pronounced for the ''angry'' models. The ''sad'' models are characterized by a relative emphasis of the lower end of the spectrum with respect to the high end. The ''admiring'' model depicts a rather flat noisy spectrum due to breathy phonation. The comparison of ''angry'' and ''frightened'' model vectors suggested that ''frightened'' samples had relatively more energy at the fundamental frequency and less energy around the main formant frequencies at 1–4 kHz.

III. DISCUSSION

A. Dimensions of auditory and acoustic differentiation

Each category included samples on which more than 70% of the listeners agreed, suggesting that acoustic cues distinguishing among 10 emotional categories can be embedded in a single word. Most ''astonished,'' ''angry,'' ''frightened,'' and ''commanding'' intentions were perceived as intended, and many of them were agreed on by over 90% of the judges. The distributions of auditory judgments showed that ''angry'' was confused with ''commanding,'' ''plead-

ing'' with ''sad,'' and ''content'' with ''admiring.'' The result suggests that the confused categories share some acoustic features not present in other connotations. ''Scornful'' intents were confused by the listeners only with emotionally rather neutral ''naming,'' suggesting that the cues of ''scorn'' are quite specific.

Most variation of mean $F0$, $F0$ range, sound pressure, and duration took place during the 200–400 ms [aa] segment of [saara]. It is thus possible that the connotations could be distinguished using this segment alone. However, a relative emphasis of [s] for ''scornful,'' lengthening of [ra] with respect to [aa] for ''frightened,'' and a rise of $F0$ at the syllable boundary for ''naming'' suggest that [s] and [ra] may also carry some specific information. To distinguish among ''command,'' ''anger,'' ''fear,'' ''sadness,'' and ''naming,'' the speakers varied the sound pressure and mean $F0$ of the [aa] segment. Stereotypic intonations of the [aa] segment, deviating from the habitual gradually falling contour, were used for ''astonished,'' ''pleading,'' and ''scornful.'' A breathy voice quality was used for ''admiring.'' The observations on the acoustic features are in line with the suggestion, made in a study on artificially modified sentenced speech, that expressions of speaker affect are due to independent variations of mean $F0$, intonation contour, and voice quality (Ladd *et al.*, 1985).

B. Significance of individual speech range

Intraspeaker comparisons of sound pressure and mean $F0$ showed that the speakers used higher $F0$ for ''frightened'' than ''angry,'' and higher sound pressure for ''angry'' than ''frightened.'' This regularity in the speakers' performance is not detected by comparing the group means because it is then obscured by interindividual variation of $F0$ and pressure range. The small changes within an individual's range may, however, contribute to auditory discrimination between these two categories.

Characteristic changes take place in the glottal waveform when the habitual speech mode is changed to ''loud'' or ''soft'' (Holmberg *et al.*, 1988; Cummings and Clements, 1995), or when the speaker is asked to express various emotions (Cummings and Clements, 1995; Laukkanen *et al.*, 1996). These glottal wave characteristics distinguish among different speech modes irrespective of interspeaker differences in $F0$ and sound pressure (Cummings and Clements, 1995; Laukkanen *et al.*, 1996). They also induce auditory cues allowing correct judgment of the speaker's effort from a recorded signal at different playback intensities (Brandt, 1972). The affect-related glottal wave characteristics contribute to correct auditory recognition of emotional connotations in the speech of subjects with different ranges of $F0$ and sound pressure, and at different distances from a (loud) speaker.

A speaker's effort to increase the sound pressure is accompanied by a rise of $F0$ (Skinner, 1935; Monsen and Engebretson, 1977; Holmberg *et al.*, 1988). In the present study, variation of sound pressure with emotional–motivational connotations suggested that at least four different effort levels were used: habitual conversation level for ''naming/astonished/pleading/admiring/content/scornful,''

somewhat lower for ''sad,'' somewhat higher for ''commanding,'' and very high for both ''angry'' and ''frightened.'' The increased effort in emotive speech was, however, not accompanied by a reflexive increase in $F0$ as during the regulation of the intensity of neutral speech: for instance, when changing from ''naming'' to ''commanding'' some speakers increased sound pressure without increasing $F0$.

C. Extraction of perceptually meaningful spectral features

The spectral models created by the self-organizing map algorithm provided information similar to the skewness of the formant region (Novák and Vokřál, 1993) or spectral slopes (van Bezooyen, 1984; Banse and Scherer, 1996) which have been used to distinguish among samples with different degrees of perceived loudness, tension (''squeezed and tight quality''), and laxness (''sonorous and resonant quality;'' van Bezooyen, 1984) and with different emotional connotations (Novák and Vokřál, 1993; Banse and Scherer, 1996).

The speech spectrum carries much more information about the glottal waveform than $F0$ and sound pressure, and cues induced by variations in the vocal tract shape with the speech mode are also embedded in it. Although possible, the determination of spectral criteria which would allow a good separation among different speech modes is a complicated task because of high intra- and interspeaker variability of the signal. To distinguish the variation due to the speech mode from other sources of variation (articulation and individual characteristics of voice and speech), it is necessary to select the individual spectral features (e.g., characteristics of spectral filters) using the error rate in the classification of perceptually categorized samples as a criterion. Statistically representative sets of perceptually categorized speech samples are needed for the optimization of feature extraction.

In the present study, the judges rated the samples by selecting one of 10 alternative categories. The distributions of the selections suggested that many samples carried ambiguous or multiple connotations. More information about what is perceived is obtained if each sample is rated simultaneously along several semantic dimensions using, for instance, 10 visual analog scales (each scale ranging from ''not at all'' to ''extremely much''). Categorization of samples with ten-dimensional evaluation patterns instead of category labels would benefit the selection of acoustic features with respect to their perceived significance. We have used such approach for the categorization of dysphonic voice qualities for automatic feature selection (Leinonen *et al.*, 1997).

D. Anger and fear

Aggression and fear have been among the best-conveyed connotations also in other studies on emotive speech (Fairbanks and Pronovost, 1939; Pfaff, 1954; Davitz, 1964; Ross *et al.*, 1973; Fenster *et al.*, 1977; van Bezooyen, 1984; Johnson *et al.*, 1986; Banse and Scherer, 1996), and singing (Kotlyar and Morozov, 1976). The ''angry, cross, furious'' and ''frightened, terrified'' frame stories of the present study pictured sudden highly arousing situations. When referring

to less panicky situations, fear may be confused with ''anxiety'' (Banse and Scherer, 1996), or ''sadness'' (Davitz and Davitz, 1959; van Bezooyen, 1984).

In the present study, both the ''angry'' and ''frightened'' connotations were expressed with high speech pressure and high mean $F0$ which is in agreement with several previous reports (Fairbanks and Pronovost, 1939; Höffe, 1960; Williams and Stevens, 1972; Banse and Scherer, 1996). The ''angry'' utterances of the present study were spoken with higher speech pressure and lower $F0$ than the ''frightened'' ones, suggesting that these measures were related to audible differences in the speech mode. A relative emphasis of the main formant frequencies in the ''angry'' utterances, suggested by the present study, has previously been described by Williams and Stevens (1972) and Novák and Vokřál (1993). The indices computed by Banse and Scherer (1996) for energy variation at different spectral bands also suggest that the proportion of spectral energy at 1–4 kHz is higher for ''angry'' than ''frightened'' samples.

In the present study, ''frightened'' was conveyed better in the female than male utterances. In a study on 30 adult male and 30 male child speakers, it was found that anger was well-recognized in both adult and child speech but fear was recognized in child speech only (Fenster et al., 1977). Such observations suggest that men are poor in the vocal expression of fear. Such intergender differences are to be expected because of sexual differentiation of both the vocal apparatus and the brain. Sexual differentiation of vocal behavior has been documented for many primate species (Mitchell, 1979; Waser, 1982).

E. Astonishment

In the present study, all the ''astonished'' intents were correctly conveyed. The single word carrier was probably more suited than connected speech for the expression of the feeling elicited by the frame story. In connected speech, the conveyance of surprise/astonishment is often somewhat poorer than that of anger and fear (Davitz, 1964; Ross et al., 1973). The good conveyance of ''astonishment'' was probably due to the stereotypic intonation with an initially high $F0$, followed by a rapid fall. The finding is in agreement with other observations on surprised speech: a wide pitch range (Fónagy and Bérard, 1972; van Bezooyen, 1984; Laukkanen et al., 1996), and a sudden rise followed by a fall (Fónagy and Magdics, 1963).

The good conveyance of ''astonishment'' suggests that signals of ''astonishment'' are of high communicative significance. The expression of surprise is often confused with that of interest (van Bezooyen, 1984) which is also characterized by an increase in the pitch range (Huttar, 1968; van Bezooyen, 1984). The $F0$ contours of [ja] (yes in German) uttered with cheerful excitement (Höffe, 1969) also resemble those of our ''astonished'' samples. It may thus be suggested that signals of positive astonishment function as signs of interest and positive excitement in everyday conversation.

F. Sadness and plea

Sadness in speech (Fenster et al., 1977; van Bezooyen, 1984; Johnson et al., 1986; Banse and Scherer, 1996) and singing (Kotlyar and Morozov, 1976) is usually well-conveyed. The present results are in agreement with previous reports indicating that in comparison to emotionally neutral speech (here represented by ''naming'') sadness is characterized by a small lengthening of syllable duration (Kotlyar and Morozov, 1976; Banse and Scherer, 1996), a decrease in intensity (Skinner, 1935; Banse and Scherer, 1996), attenuation of formant frequencies (Skinner, 1935; Kaiser, 1962; Williams and Stevens, 1972; Novák and Vokral, 1993; Banse and Scherer, 1996), and a slight fall in $F0$ (Sedláček and Sychra, 1963; Williams and Stevens, 1972; Banse and Scherer, 1996).

''Sadness'' may be confused with ''shame'' (van Bezooyen, 1984) and ''despair'' (Banse and Scherer, 1996). In the present study it was confused with ''pleading.'' The exceptional falling–rising intonation at the first syllables of the best-conveyed ''pleading'' samples suggests that the intonation is critical for the differentiation between ''sad'' and ''pleading.'' A narrow pitch range and a rising intonation on the stressed syllables have also been described for ''longing'' intents in Hungarian by Fónagy and Magdics (1963). They also reported that longing is expressed in music by slow tempo and a melody which descends and then slightly rises at the end of the motive. These descriptions bring ''longing'' acoustically near our ''pleading'' category. Interestingly, we did not find any previous studies on human vocal expression of ''pleading/submissive.'' In man, as in the monkey, the pleading mode may be characteristic of young individuals who have been rewarded for using it by physical contact or favorite food.

G. Contentment and admiration

Our results agree with the observations that ''contentment'' and ''happiness'' (Fenster et al., 1977), or ''admiration,'' ''satisfaction,'' and ''cheerfulness'' (Davitz, 1964), are difficult to convey. In the present study, ''content'' and ''admiring'' connotations were not well-distinguished by the listeners. Compared with ''naming,'' the utterances of both intents were much longer. A similar prolongation of utterances of [ja] (yes) when changing the mode from a simple statement (sachlich) to content (hingebungsvoll, glücklich) was described by Höffe (1960). The soft breathy voice with a flat noisy spectrum in the best-conveyed ''admiring'' samples suggested that this characteristic was critical for the conveyance of ''admiration'' and for its differentiation from ''content.'' When simulating ''admiring'' and ''content'' the speakers often smiled, and the audible cues of smiling (Fónagy, 1967; Tartter, 1980; Tartter and Braun, 1994) may distinguish ''content'' and ''admiring'' from the other connotations of the study.

H. Scorn

Contempt in speech is usually well-conveyed (Fairbanks and Pronovost, 1939; Kramer, 1964; Ross et al., 1973; Banse and Scherer, 1996). In the present study, some ''scornful''

1861   J. Acoust. Soc. Am., Vol. 102, No. 3, September 1997

Leinonen et al.: Expression of emotions in speech   1861

samples were agreed on by over 80% of the judges, whereas some were not distinguished from ''naming.'' Thus not all speakers succeeded in simulating ''scorn.'' The well-conveyed samples were marked by an extreme fall of $F0$ (down to vocal fry) which is in agreement with previous observations (Fairbanks and Pronovost, 1939; Fónagy and Magdics, 1963; Banse and Scherer, 1996). Our results also agree with descriptions on lengthening of stressed syllables (Fónagy and Magdics, 1963) and prolongation of phonation duration (Fairbanks and Hoagling, 1941) when changing from neutral to scorn/contempt.

I. Naming and command

In our study, the ''naming'' (i.e., ''naming, informing, addressing'') category was planned to represent the ''neutral'' mode. Unexpectedly, most utterances with ''naming'' intents had a higher mean $F0$ at the second than first syllable, which is against the Finnish norm for neutral connected speech. In the auditory rating test, however, samples with both rising and falling word intonation were categorized as ''naming.'' The rising intonation was thus not critical for the ''naming'' connotation. Possibly, it signaled something else about the speaker of the frame story: the chairman of a meeting uttering Saara's name as an invitation to speak, or as a question prompting an answer.

The ''naming'' intents contrasted with ''commanding'' which always had a regular and pronounced fall of mean $F0$ at the second syllable. The results suggest that the commanding mode can be avoided by raising the pitch at the end of the utterance. A study on perceptual differentiation of different artificial intonation contours has led to a similar suggestion: Authoritative connotation relates to a tendency to a final fall, whereas submissive to a final rise (Uldall, 1960).

The recognition of the neutral/indifferent connotation among several emotional connotations is usually accurate (Fairbanks and Pronovost, 1939; Ross *et al.*, 1973; van Bezooyen, 1984). A similar observation was made on the ''naming'' connotation. As reviewed in detail above (in connection with the other connotations), the ''naming'' samples resembled the neutral/indifferent mode of other studies in the following respects: the phonation duration was somewhat shorter than in ''sad,'' ''content,'' ''admiring,'' ''scornful,'' and ''astonished,'' and the $F0$ level was higher than in ''sad'' or ''contempt.'' The change from ''naming'' to ''commanding'' mode in the present study was similar to that described by Kaiser (1962) for the change from ''neutral'' to ''grim'': relative enhancement of main formant frequencies, steepening of the falling intonation, and some shortening of duration. Such cues may account for listeners' agreement about ''leadership'' in speech samples (Fay and Middleton, 1941).

IV. CONCLUSION

All 10 connotations expressed with the word [saara] were distinguished by the listeners. The acoustic analyses suggest that connotations ''naming,'' ''sad,'' ''commanding,'' ''angry,'' and ''frightened'' were distinguished using cues which were partly related to the variation of mean $F0$

and peak sound pressure within the individual's range. The recognition of ''pleading,'' ''admiring,'' ''content,'' ''astonished,'' and ''scornful'' was made possible by specific patterns of intonation and specific voice qualities not related to variation of mean $F0$ and sound pressure. Comparison with studies conducted on other European languages suggests that the speakers use similar means to express emotional neutrality, scorn, positive astonishment, fear, anger, and sadness; there is not enough data for the evaluation of intercultural communication of contentment, admiration, dominance, and plea/submissiveness.

Banse, R., and Scherer, K. R. (**1996**). ''Acoustic profiles in vocal emotion expression,'' J. Pers. Soc. Psychol. **70**, 614–636.

Bezooyen, R. van (**1984**). *Characteristics and Recognizability of Vocal Expressions of Emotion* (Foris, Dordrecht), pp. 1–165.

Brandt, J. F. (**1972**). ''Effects of stimulus bandwidth on listener judgment of vocal loudness and effort,'' J. Acoust. Soc. Am. **52**, 705–707.

Cummings, K. E., and Clemens, M. A. (**1995**). ''Analysis of the glottal excitation of emotionally styled and stressed speech,'' J. Acoust. Soc. Am. **98**, 88–98.

Davitz, J. R., and Davitz, L. J. (**1959**). ''The communication of feelings by content-free speech,'' J. Commun. **9**, 6–13.

Davitz, J. R. (**1964**). ''Auditory correlates of vocal expressions of emotional feelings,'' in *The Communication of Emotional Meaning*, edited by J. R. Davitz (McGraw-Hill, New York), pp. 101–112.

Fairbanks, G., and Hoaglin, L. W. (**1941**). ''An experimental study of the durational characteristics of the voice during the expression of emotion,'' Speech Monogr. **8**, 85–90.

Fairbanks, G., and Pronovost, W. (**1939**). ''An experimental study of the pitch characteristics of the voice during the expression of emotion,'' Speech Monogr. **6**, 87–104.

Fay, P. J., and Middleton, W. C. (**1943**). ''Judgment of leadership from the transmitted voice,'' J. Soc. Psychol. **17**, 99–102.

Fenster, C. A., Blake, L. K., and Goldstein, A. M. (**1977**). ''Accuracy of vocal emotional communications among children and adults and the power of negative emotions,'' J. Commun. Disord. **10**, 301–314.

Fónagy, I. (**1967**). ''Hörbare Mimik,'' Phonetica **16**, 25–35.

Fónagy, I., and Bérard, E. (**1972**). ''Il est huit heures: Contribution à l'analyse sémantique de la vive voix,'' Phonetica **26**, 157–192.

Fónagy, I., and Magdics, K. (**1963**). ''Emotional patterns in intonation and music,'' Z. Phonet. Sprachwissen. Kommun. **16**, 293–326.

Holmberg, E. G., 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**, 511–527.

Huttar, G. L. (**1968**). ''Relations between prosodic variables and emotions in normal American English utterances,'' J. Speech Hear. Res. **11**, 467–480.

Höffe, W. L. (**1960**). ''Über Beziehungen von Sprachmelodie und Lautstärke,'' Phonetica **5**, 129–159.

Johnson, W. F., Emde, R. N., Scherer, K. R., and Klinnert, M. D. (**1986**). ''Recognition of emotion from vocal cues,'' Arch. Gen. Psychiat. **43**, 280–283.

Kaiser, L. (**1962**). ''Communication of affect by single vowels,'' Synthese **14**, 300–319.

Karlsson, F. (**1987**). *Finnish Grammar* (Werner Söderström, Porvoo), pp. 18, 20.

Kohonen, T. (**1995**). *Self-Organizing Maps* (Springer, Berlin), pp. 77–127.

Kohonen, T., Kangas, J., and Laaksonen, J. (**1992**). *SOM_PAK, the Self-Organizing Map Program Package*. Available via anonymous ftp at the Internet site cochlea.hut.fi.

Kotlyar, G. M., and Morozov, V. P. (**1976**). ''Acoustic correlates of the emotional content of vocalized speech,'' Sov. Phys. Acoust. **22**, 208–211.

Kramer, E. (**1964**). ''Elimination of verbal cues in judgments of emotion from voice,'' J. Abn. Soc. Psychol. **68**, 390–396.

Ladd, D. R., Silverman, K. E. A., Tolkmitt, F., Bergmann, G., and Scherer, K. R. (**1985**). ''Evidence for the independent function of intonation contour type, voice quality, and $F0$ range in signaling speaker affect,'' J. Acoust. Soc. Am. **78**, 435–444.

Laukkanen, A.-M., Vilkman, E., Alku, P., and Oksanen, H. (**1996**). ''Physical variations related to stress and emotional state: A preliminary study,'' J. Phonetics **24**, 313–335.

Leinonen, L., Linnankoski, I., Laakso, M.-L., and Aulanko, R. (**1991**). ''Vocal communication between species: man and macaque,'' Lang. Commun. **11**, 241–262.

Leinonen, L., Hiltunen, T., Torkkola, K., and Kangas, J. (**1993**). ''Self-organized acoustic feature map in detection of fricative-vowel coarticulation,'' J. Acoust. Soc. Am. **93**, 3468–3474.

Leinonen, L., Hiltunen, T., Laakso, M-L., Rihkanen, H., and Poppius, H. (**1997**). ''Categorization of voice disorders with six perceptual dimensions,'' Folia Phoniat. Logop. **49**, 9–20.

Linnankoski, I., Laakso, M., Aulanko, R., and Leinonen, L. (**1994**). ''Recognition of emotion in macaque vocalizations by children and adults,'' Lang. Commun. **14**, 183–192.

McCluskey, K. W., Albas, D. C., Niemi, R. R., Cuevas, C., and Ferrer, C. A. (**1975**). ''Cross-cultural differences in the perception of the emotional content of speech: a study of the development of sensitivity in Canadian and Mexican children,'' Dev. Psychol. **11**, 551–555.

Mitchell, G. (**1979**). *Behavioral Sex Differences in Nonhuman Primates* (Van Nostrand Reinhold, New York), pp. 335–355.

Monsen, R. B., and Engebretson, A. M. (**1977**). ''Study of variations in the male and female glottal wave,'' J. Acoust. Soc. Am. **62**, 981–993.

Murray, I. R., and Arnott, J. L. (**1993**). ''Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion,'' J. Acoust. Soc. Am. **93**, 1097–1108.

Novák, A., and Vokřál, J. (**1993**). ''Emotions in the sight of long-time averaged spectrum and three-dimensional analysis of periodicity,'' Folia Phoniat. **45**, 198–203.

Pfaff, P. L. (**1954**). ''An experimental study of the communication of feeling without contextual material,'' Speech Monogr. **21**, 155–156.

Rihkanen, H., Leinonen, L., Hiltunen, T., and Kangas, J. (**1994**). ''Spectral pattern recognition of improved voice quality,'' J. Voice **8**, 320–326.

Ross, M., Duffy, R. J., Cooker, H. S., and Sargeant, R. L. (**1973**). ''Contribution of the lower audible frequencies to the recognition of emotions,'' Am. Ann. Deaf **118**, 37–42.

Sammon, J. W., Jr. (**1969**). ''A non-linear mapping for data structure analysis,'' IEEE Trans. Comp. **C-18**, 401–409.

Scherer, K. R. (**1986**). ''Vocal affect expression: a review and a model for future research,'' Psychol. Bull. **99**, 143–165.

Scherer, K. R. (**1995**). ''Expression of emotion in voice and music,'' J. Voice **9**, 235–248.

Sedláček, K. von, and Sychra, A. (**1963**). ''Die Melodie als Faktor des emotionellen Ausdrucks,'' Folia Phoniat. **15**, 89–98.

Skinner, E. R. (**1935**). ''A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness,'' Speech Monogr. **2**, 81–137.

Tartter, V. C. (**1980**). ''Happy talk: perceptual and acoustic effects of smiling on speech,'' Percept. Psychophys. **27**, 24–27.

Tartter, V. C., and Braun, D. (**1994**). ''Hearing smiles and frowns in normal and whisper registers,'' J. Acoust. Soc. Am. **96**, 2101–2107.

Uldall, E. (**1960**). ''Attitudinal meanings conveyed by intonation contours,'' Lang. Speech **3**, 223–234.

Waser, P. M. (**1982**). ''The evolution of male loud calls among mangabeys and baboons,'' in *Primate Communication*, edited by C. T. Snowdon, C. H. Brown, and M. R. Petersen (Cambridge U.P., Cambridge), pp. 117–143.

Williams, C. E., and Stevens, K. N. (**1972**). ''Emotions and speech: some acoustical correlates,'' J. Acoust. Soc. Am. **52**, 1238–1250.

Speaker normalization in the perception of Mandarin Chinese tones

Corinne B. Moore[a] and Allard Jongman[b]

Cornell Phonetics Laboratory, Department of Modern Languages, Morrill Hall, Cornell University, Ithaca, New York 14853

This study investigated speaker normalization in perception of Mandarin tone 2 (midrising) and tone 3 (low-falling–rising) by examining listeners' use of $F0$ range as a cue to speaker identity. Two speakers were selected such that tone 2 of the low-pitched speaker and tone 3 of the high-pitched speaker occurred at equivalent $F0$ heights. Production and perception experiments determined that turning point (or inflection point of the tone), and $\Delta F0$ (the difference in $F0$ between onset and turning point) distinguished the two tones. Three tone continua varying in either turning point, $\Delta F0$, or both acoustic dimensions, were then appended to a natural precursor phrase from each of the two speakers. Results showed identification shifts such that identical stimuli were identified as low tones for the high precursor condition, but as high tones for the low precursor condition. Stimuli varying in turning point showed no significant shift, suggesting that listeners normalize only when the precursor varies in the same dimension as the stimuli. The magnitude of the shift was greater for stimuli varying only in $\Delta F0$, as compared to stimuli varying in both turning point and $\Delta F0$, indicating that normalization effects are reduced for stimuli more closely matching natural speech. © *1997 Acoustical Society of America.* [S0001-4966(97)01408-2]

PACS numbers: 43.71.Bp, 43.71.Es [WS]

INTRODUCTION

This study examines the role of speaker-dependent $F0$ information in the perception of lexical tone. It is well known that the perception of segments requires listeners to normalize to reduce overlap among phonetic categories. A classic example of this overlap is illustrated by formant frequency data for vowels in which two phonetic categories from different speakers have similar formant values (Peterson and Barney, 1952). Listeners either actively or passively compensate for this acoustic variability, caused by differences in speaker vocal tract size, in order to identify segments accurately. Consequently, segments that have very similar acoustic characteristics may not be perceived identically.

A. Speaker normalization: The use of speaker-specific acoustic information in vowel perception

Previous work on normalization has shown that listeners use acoustic information outside of the speech sound itself (extrinsic information) about speaker identity in order to classify vowels. For example, Peterson and Barney (1952) found that perception of vowel tokens produced by a wide variety of speakers exhibited confusion within the areas of overlap in the vowel formant data. Ladefoged and Broadbent (1957) provided evidence that listeners refer to extrinsic information specifying the vowel space of a speaker. In this study, six versions of the phrase *please say what this word is* were synthesized, each with a different range of $F1$ and $F2$ to represent different speakers. In addition, four test words of

the form, ''*b_t*'' were synthesized with $F1$ and $F2$ values of the vowel approximately corresponding to the vowels in the words ''bit,'' ''bet,'' ''bat,'' and ''but.'' Listeners were asked to identify the final target word. Results showed that perception of a given target word changed as a function of the formant frequency range in the precursor. For example, the target *bit* would be perceived as *bet* when preceded by the precursor in which $F1$ was relatively low. There was thus a contrast effect, whereby a low $F1$ in the precursor would make listeners perceive a given $F1$ in the target as relatively higher, changing the percept from [ɪ] to [ɛ].

While Ladefoged and Broadbent's results suggested *what* acoustic information was used in speaker normalization, it remained unclear *how* this information was used. In particular, acoustic information may either be used to identify the speech sound directly, or it may be used as a cue to speaker identity, establishing a representation against which acoustic characteristics may be calibrated. This question was investigated by Johnson (1990) for vowel continua. In Johnson's study, speaker identity was defined by $F0$, whereas Ladefoged and Broadbent (1957) manipulated formant frequencies to specify speaker identity.

In a series of three experiments, Johnson examined perception of test words in isolation and in carrier phrases whose $F0$ manipulations signaled the same speaker, different speakers, or were ambiguous with respect to speaker identity. In a series of perception pretests, Johnson manipulated $F0$ in a synthesized ''hood-hud'' continuum and also in the synthesized carrier phrase ''this is—.'' These pretests were designed to determine the relationship between $F0$ and speaker identity for both the vowel tokens and the carrier phrases. In the first experiment, Johnson compared percep-

[a]Current address: Corporate Technical Publications, Diebold Incorporated, 5995 Mayfair Road, North Canton, OH 44720.
[b]Electronic mail: aj12@cornell.edu

tion of vowels in isolation and in carrier phrases. The $F0$ levels of the vowel tokens were 100 and 150 Hz, levels which had been shown by the pretests to correspond to different speakers. These vowels were also embedded in carrier phrases which had been attributed to a single speaker. Listeners were asked to label the vowel tokens. The results of the vowels in carrier phrases were then compared to results of those vowels in isolation. Results showed that shifts in vowel identification observed for the 150-Hz versus the 100-Hz tokens were reduced if the carrier phrase signaled that they were produced by the same speaker. The second experiment compared the vowel shifts when both vowel tokens and carrier phrases were ambiguous with respect to whether they had been produced by the same or different speakers. Results showed no significant difference in shifts for vowel tokens in isolation as compared to those in the carrier phrases, presumably since there was no conflicting information about the speaker. Finally, Johnson showed that perception of vowels shifted when the carrier phrases indicated two speakers, but the vowel tokens were at a constant $F0$ level, corresponding to a single speaker. This result supported Ladefoged and Broadbent (1957), in that identical stimuli were perceived differently when perceived as being produced by different speakers. The three experiments in Johnson's study thus provided evidence that listeners use $F0$ as a cue to speaker identity, and that this perceived speaker identity affects vowel perception.

B. Context effects and Mandarin Chinese tones

While the majority of studies have investigated normalization in the perception of vowels, less experimental work has been done to examine speaker normalization in the perception of other types of speech sounds (for examples of normalization for consonants, see Mann and Repp, 1980; Whalen, 1981; Jongman and Miller, 1991; Johnson, 1991). The present study extends work on normalization from the segmental domain to the suprasegmental domain, focusing on lexical tone. The tone language used in this study is Mandarin Chinese, whose four lexical tones include a high-level tone (tone 1), a midrising tone (tone 2), a low-falling–rising tone (tone 3), and a high-falling tone (tone 4). Examples of the four tones for one speaker are shown in Fig. 1. The present series of experiments examines whether $F0$ range, as a cue to speaker identity, influences perception of tones 2 and 3.

As suprasegmentals, tones are perceived relative to other tones, although they are also distinguished by tone-internal (intrinsic) acoustic properties—primarily pitch height and contour (Gandour, 1978; Coster and Kratochvil, 1984). For tones which contrast in both of these dimensions, intrinsic $F0$ information may be sufficient for correct identification. To identify tones differing only in $F0$ height, however, listeners must refer to their knowledge of the speaker's $F0$ range, and where tones occur within that range. For example, a low tone produced by a high-pitched speaker and a high tone produced by a low-pitched speaker may be acoustically very similar. The process by which listeners adjust perception according to speaker-specific acoustic information is referred to as speaker normalization. Few studies have inves-
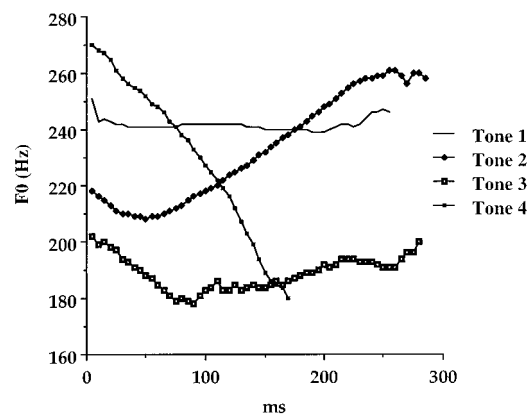
FIG. 1. $F0$ contours for each of the four Mandarin Chinese tones, taken from one token spoken in isolation by one of the speakers in this study (segmental context *ma*).

tigated the role of extrinsic $F0$ in tone perception, however, and results from these studies have not provided convincing evidence for speaker normalization.

In a study specifically addressing speaker normalization for tones, Leather (1983) tested perception of Mandarin Chinese tone stimuli in two natural precursor phrases, one representing a low $F0$ range, the other one a higher range. Seven stimuli from two tone 1-tone 2 continua, each continuum representing the $F0$ range of one speaker, were embedded in the precursor phrases. Four steps in the middle of the continuum were identical in $F0$ height and contour, and were included in both continua. Test items (precursor+tone stimulus) were blocked by speaker and presented to five listeners in a labeling task. Individual subject responses were reported for the four pairs of midcontinuum stimuli (paired by speaker condition). Results showed that perception of at least one stimulus pair varied as a function of the speaker.

Unfortunately, Leather did not explicitly predict the type of responses he expected in each condition, nor did the reported data present this information. It is difficult, therefore, to take these data as conclusive evidence for speaker normalization without more detailed information. In particular, it is essential to know the direction of any shift in identification, whether it is consistent across speakers and across stimuli, and whether subject responses conform to predicted results. Leather's use of the tone 1-tone 2 continuum may also have been problematic. These two tones, which vary in both $F0$ height and contour, were synthesized without controlling for confounding acoustic parameters such as onset or offset $F0$. Moreover, both tones occur in the upper region of a speaker's pitch range, making it difficult to compare these tones in terms of $F0$ height.

Other studies have examined the role of extrinsic $F0$ information in tone perception, though they did not specifically address speaker normalization. Using an AX anchoring paradigm, in which the A element of the stimuli was constant, Lin and Wang (1985) presented subjects with pairs of Mandarin Chinese tones in which the first tone, representing a high-level tone (tone 1), was held at a constant 115 Hz, while the second tone, representing the high-falling tone (tone 4), varied onset $F0$ from 110 to 140 Hz in 10 Hz steps

with an $F0$ fall of 40 Hz. Subjects were asked to label the first tone in each pair. Their results showed that as the onset $F0$ in the second syllable increased, identification of the first tone as a rising tone (tone 2) increased. Thus the higher onset $F0$ of the second syllable cued a wider pitch range, altering the relative $F0$ height of the first tone 1 syllable to be perceived as low. Without a statistical analysis it is uncertain how robust these results are, but they nevertheless provide some evidence that tones are perceived relative to $F0$ range, such that this information contributes directly to the acoustic characteristics of the tone. While the study more broadly indicates that tone perception is affected by extrinsic $F0$, it does not address whether $F0$ information which serves to distinguish speaker identity may influence perception.

Using a similar anchoring paradigm, Fox and Qi (1990) investigated whether context $F0$ influences tone perception, and whether the influence occurs for both native and non-native listeners. Tone stimuli were presented in isolation and in pairs. In the isolated-token condition, listeners were asked to rate the stimulus according to how closely it resembled the tone 1 or tone 2 exemplar. In the paired-token condition, the first tone was either a tone 1 or tone 2, while the onset $F0$ of the second tone varied along a continuum from tones 1 to 2; subjects were asked to rate the second tone in the pair, according to the same rating scale as in the isolated-token condition. Results showed no significant difference between perception in isolation and in the context condition for either language group.

Following Leather's study, Fox and Qi (1990) presented chi-square values for individual subject responses to four midcontinuum stimuli, showing highly inconsistent patterns of identification across subjects and stimuli. Fox and Qi interpreted these results as weak support for context effects from $F0$ on tone perception, in contrast to those of Lin and Wang (1985), who showed differences in identification as $F0$ range widened.

The reasons for the inconsistencies in Fox and Qi may be related to the methodology used. In Lin and Wang (1985), manipulating the onset of the second tone had the effect of modifying the pitch range, as in Fox and Qi, but listeners were asked to identify the first tone in the sequence, a tone which was constant throughout the experiment. In comparison, Fox and Qi asked listeners to identify the tone containing the modifications, the second tone. The anchor in Fox and Qi did not shift, but rather it was intended that listeners would use the anchor to identify the onset of the second tone as lower, as in a tone 2, or higher, corresponding to a tone 1. A shift in identification for tone 1 anchors may have been expected, since listeners may not have had enough $F0$ range information against which to calibrate the tone stimuli. However, a tone 2 anchor would provide the listener with adequate pitch range information against which to compare the $F0$ onset of the second tone. Since both anchors were included in one test, listeners had the relevant $F0$ range information throughout the test. Therefore it is not surprising that the results yielded no context effects.

Results from these earlier studies have not provided robust evidence that tone perception is affected by contextual acoustic cues, despite the assumption that tones, as suprasegmentals, are perceived according to surrounding information. In Leather 1983 as well as Fox and Qi (1990), shifts in tone identification did not occur reliably for all subjects, nor for a particular stimulus. Also, the direction of the shift, whether contrastive or assimilatory, was either not specified or was inconsistent across subjects and stimuli.

Some of these problems may be remedied by employing a different methodology. For example, in order to test for speaker normalization, precursors must vary in speaker identity. Precursors in Leather (1983) represented different speakers, but Lin and Wang (1985) and Fox and Qi (1990) limited their investigation to context effects, and so precursors consisted of one syllable which did not represent more than one speaker. Moreover, stimuli should reflect a situation in which normalization would be expected to occur, for example, in perception of different tones occurring within an area of overlap in $F0$ range among speakers. Although Leather (1983) examined tone perception for speakers with overlapping $F0$ ranges, both of the tones occurring in that range were high tones (tones 1 and 2), and so may not have been sufficiently distinguished by $F0$ height. In addition, subject data should be analyzed over the entire continuum, rather than for selected stimuli, to determine whether the identification functions for each subject have shifted reliably, and in what direction. The present study addresses these issues in a new investigation of speaker normalization for Mandarin Chinese tones.

In this study, production and perception tests were used to examine tone 2 and tone 3. These tones were chosen, as opposed to tones 1 and 2 used by Leather (1983) and Fox and Qi (1990), because they occupy distinct registers in a speaker's range; although both tones originate at the midpoint of the range, tone 2 rises to cover the high region of the range, while tone 3 is distinctly low, falling to the low region and ending with a rise (in prepausal position) near the middle of the range. This distinction more clearly demarcates $F0$ height as a perceptual cue. Tones 2 and 3 are also similar in contour when spoken in isolation, which may be the reason they cause the most confusion in perception tests (Kiriloff, 1969; Chuang *et al.*, 1972; Gandour, 1978; Li and Thompson, 1978). While overall $F0$ height may contribute to the distinctive phonetic characteristics of tones 2 and 3, two additional acoustic dimensions are relevant: Timing of the turning point, defined as the duration from the onset of the tone to the point of change in $F0$ direction, and also the decrease in $F0$ from the onset of the tone to the turning point, hereafter called $\Delta F0$. These properties are schematized in Fig. 2. Perception studies of Mandarin tones 2 and 3 have found that both timing of the turning point and $\Delta F0$ are perceptually relevant for identification of the tones (Shen and Lin, 1991; Shen *et al.*, 1993).[1]

Using these two acoustic dimensions, the present experiment examined perception of stimuli in a tone 2-3 continuum whose $F0$ levels fall within an area of overlap in $F0$ range for two speakers. In this scenario, speakers overlap in $F0$ range such that the low region of the high-pitched speaker overlaps with the high region of the low-pitched speaker. Within the area of overlap, tones may occur at equivalent $F0$ heights such that they would be low tones for the high-
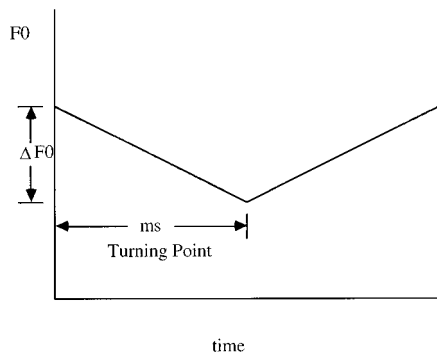
FIG. 2. Turning point and $\Delta F0$ properties schematized for a contour tone.

pitched speaker, but high tones for the low-pitched speaker. Identification of the tone would then be expected to shift (contrastively) depending on the $F0$ range of the precursor. If normalization occurs, stimuli will be identified by using $F0$ range information to ''calibrate'' ambiguous tones. On the other hand, if tone identification does not shift as a function of different $F0$ ranges, speaker normalization will be judged not to have occurred (subjects will not have referred to talker $F0$ range in order to identify tones).

To achieve the scenario conducive to normalization, production data were gathered in order to find two speakers whose $F0$ ranges and tones exhibited areas of overlap. Data from the production study also provided acoustic measurements of tones 2 and 3, which were then used in synthesizing stimuli for the perception experiments.

Stimuli for the perception tests were synthesized to vary in either $\Delta F0$, timing of the turning point, or both acoustic dimensions. These stimuli were presented to listeners in isolation, and then embedded in both high and low precursor phrases. Perception of stimuli in these two conditions was compared to determine whether changes in perceived speaker identity produce changes in tone identification.

I. EXPERIMENT 1: PRODUCTION

This experiment was designed to provide acoustic information about speaker $F0$ ranges and Mandarin tones 2 and 3. The experiment consisted of three reading tasks, the results of which established the mean $F0$ and overall $F0$ range of the speakers.

A. Method

1. Subjects

Four female and three male subjects aged between 19 and 30 years produced the data for this study. The subjects were all native speakers of Mandarin Chinese from Mainland China. They all were graduate or undergraduate students at Cornell University, and competent English speakers as well. None reported any speech disorders. Subjects were paid for their participation.

2. Materials

The data collected were from three reading tasks. The first of these asked subjects to read a long passage from a story, approximately four minutes long, entitled ''*Guo ji da*

shi he ta de qi zi'' ''The World Master and His Wife'' by Xiao Fu Xin (Hsu, 1990). For the second task, subjects read minimal sets for each of the four Mandarin tones of the segmental contexts *wu*, *yi*, *bi*, and *ma*. These syllables were randomized and produced in the carrier phrase *Zhe ge zi nian*—(''This word is—''). The third task consisted of subjects reading a randomized list of the minimal sets spoken in isolation. Test items in the carrier phrases and in isolation were produced three times. Both lists also included fillers at the beginning and end of every page to avoid list effects. All reading materials were presented to subjects in Chinese characters.

3. Procedure and analysis

Subjects read the materials in an IAC soundproof booth. They were recorded using an Electrovoice RE20 cardioid microphone and a Carver TD-1700 cassette recorder in the Cornell Phonetics Laboratory. The data were digitized on a Sun Sparcstation 2 computer using a sampling rate of 11 kHz with 16-bit resolution, and were analyzed using Entropics WAVES+/ESPS speech analysis software.

Mean $F0$ and overall $F0$ range were obtained from computer measurements of $F0$ over the long passage. A computer program sampled $F0$ every 5 ms, then filtered out $F0$ values corresponding to a probability of voicing of less than 99%. Mean and modal $F0$ values were then calculated for each speaker.

$F0$ measurements for the minimal sets in isolation and in carrier phrases, as well as the carriers themselves, were taken every 5 ms. Average $F0$ as well as peak and valley $F0$ values for the carriers were calculated for voiced portions. Valley $F0$ values were taken to be the lowest $F0$ value in the tone, peak $F0$ the highest value. $F0$ for tones was measured beginning with the onset of the vowel, or at the first full period of the vowel if the onset of voicing resulted in ''artifact'' $F0$ values which did not appear to be congruous with following $F0$ points. Ending $F0$ values were determined to occur at the offset of voicing (probability of voicing below 99%), or at the offset of the vowel (according to the waveform and spectrogram analysis) if the data showed $F0$ values inconsistent with the path of the tone to that point. Vowel and tone duration was measured from onset to offset of periodicity in the waveform in the *yi* and *wu* syllable types, from the onset of the vowel to its offset as determined by the waveform in the *bi* and *ma* syllables. Spectrograms provided additional help in locating vowel onset and offsets, where vowel onset was marked as the onset of $F1$, and the offset of $F2$ was taken to be vowel offset.

Two acoustic dimensions were measured in isolated tones and in tones embedded in the carriers: timing of the turning point, and the change in $F0$ from the onset of the tone to the turning point ($\Delta F0$). Turning point was defined from the pitch track as the duration between the onset of the tone and the point at which the tone changed $F0$ direction from falling to rising—this point was also the valley for both tones. $\Delta F0$ was calculated to be the difference in $F0$ between the onset of the tone and the turning point. Values were averaged for all instances of the tones, as well as for each syllable type. Tokens repeated in isolation and carrier
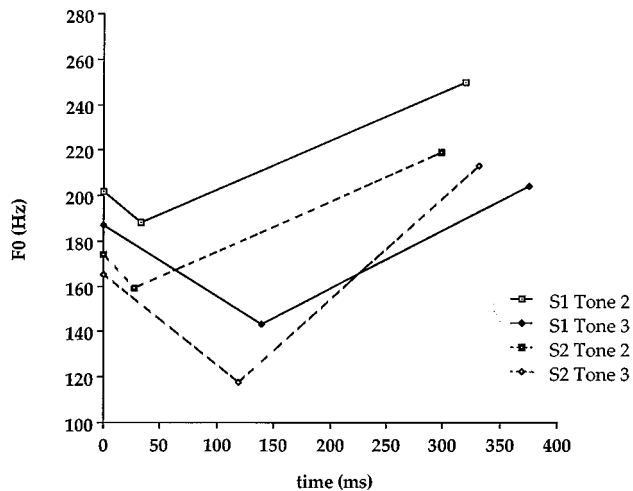
FIG. 3. Averaged $F0$ contours of tones 2 and 3 for the two female speakers (S1, solid lines; S2, dashed lines), across all syllable types.

TABLE I. Average duration (ms) of tone 2 for S2 and tone 3 for S1, according to syllable type.

Tone 2 (S2)	ma ''hemp''	yi ''move''	bi ''nose''	wu ''not''
mean	296	304	287	307
Tone 3 (S1)	ma ''horse''	yi ''to lean against''	bi ''pen''	wu ''dance''
mean	338	417	328	418

phrases comprised a corpus of 24 instances of each tone per speaker (4 syllable types×2 reading tasks×3 repetitions). Several instances of the tones contained creak, including four tone 2 and eight tone 3 tokens, which made the relevant measurements impossible, and so these tokens were excluded from analysis.

B. Results

1. F0 range data

Analysis of $F0$ range data was conducted for the long passage, generating roughly 40 000 data points for each speaker. Among seven speakers analyzed, $F0$ range data for two of the four female speakers were found to meet the requirements of the normalization study. Speaker 1 (hereafter S1) had a mean $F0$ of 212 Hz while Speaker 2 (hereafter S2), had a mean $F0$ of 192 Hz. The means reflect that S1 produced more consistently in a higher range than S2. These two female speakers, then, illustrate the $F0$ range characteristics most conducive to testing for speaker normalization. The data show a region of overlap in the $F0$ ranges for S1 and S2. Tones which occur in the overlapping region could conceivably fall in the low region of S1's range, but the high region of S2's range. The tones corresponding to those areas of the speaker ranges are tone 2, the midrising tone, which typically occurs in the upper region of a speaker's range, and tone 3, the low-falling–rising tone, which occupies the low region of a speaker's $F0$ range. If those tones are to be perceived correctly, listeners must adjust tone perception according to which speaker produced the tone.

2. Mandarin tone 2 and tone 3 analysis

Figure 3 shows the $F0$ contours of tones 2 and 3 for both speakers. The $F0$ contours in Fig. 3 represent tone 2 and tone 3 average $F0$ onset, turning point and offset values for all syllable types in the isolation and carrier conditions. Figure 3 shows that the two tones are similar in $F0$ at onset, and have a similar falling–rising contour. Importantly, the $F0$ height of S1's tone 3 falls somewhat below the tone 2 of S2 in a relationship corresponding to the overlap in $F0$

ranges; the tone 2 contour of S2 has an onset of 174 Hz, falling to 159 Hz at the turning point, and ending at 219 Hz, the tone 3 of S1 has an onset of 187 Hz, falling to 143 Hz, and ending at 204 Hz. Tone 2 for S1 is produced outside of the region of tonal overlap; although tone 3 for S1 and S2 overlap, this overlap is irrelevant since it does not involve a change in tonal category. The crucial observation is that S1's tone 3 and S2's tone 2 overlap.

Recall that the $F0$ range data indicated that the two female speakers shared a region of overlap which encompasses the lower range of S1 and the upper range of S2. The low tone of S1 and the high tone of S2 occurred precisely in this region, a pattern that is predicted by the $F0$ range data. These data attest to the likelihood that tone 2 syllables for S2 and tone 3 syllables for S1 may be confusable in isolation. They were, thus, appropriate to use in subsequent tests for normalization.

Mean duration measurements taken for each tone according to the vowel in each syllable type are shown in Table I. Table I lists average tone durations for each syllable type, including six tokens of each type (three tokens produced in isolation, three in carrier sentences), for a total of 24 tokens possible. Fourteen tokens of S1 (five for *ma*, two for *yi*, five for *bi*, and two for *wu*) and two tokens of S2's *wu* were excluded because the presence of creak made location of vowel offset impossible.[2] S2's tone 2 durations range from 268 to 371 ms; S1's tone 3 range was from 328 to 483 ms. Average duration for the tone 2 (S2) tokens was 299 ms, as compared with 375 ms for tone 3 (S1) tokens. An unpaired, two-tailed t test showed this difference to be significant $[t(30) = 7.98, p < 0.001]$. However, a substantial area of overlap was represented in these tone 2 and 3 tokens: from 328 to 371 ms.

Duration was also measured in terms of timing of the turning point for tone 2 (S2) and tone 3 (S1), over all syllable types. These data were analyzed in two ways: in absolute duration (ms), and also as a percentage of tone duration. Tone 2 turning point values averaged 27 ms, occurring at an average of 9% into the tone. Tone 3 showed an average turning point of 139 ms, occurring at 37% into the tone on average. Turning point values for tone 2 ranged from 0 to 91 ms, as compared to 105 to 200 ms for tone 3, demonstrating a significant difference between the two tones $[t(30) = 11.32, p < 0.0001]$. Calculated as a percentage of total tone duration, tone 2 ranged from 0% to 30% of the tone, while the tone 3 range was from 28% to 53%. There was, thus, a small area of overlap between the tones, although the difference between the two tones was significant $[t(30) = 13.16, p<0.0001]$.

In addition to turning point, the other acoustic parameter observed was the decline in $F0$ from the onset of the tone to

the turning point, or $\Delta F0$. These data showed an average $\Delta F0$ of 15 Hz for tone 2 (S2), and 44 Hz for tone 3 (S1). For tone 2, $\Delta F0$ ranged from 0 to 56 Hz, and from 24 to 106 Hz for tone 3. There was, thus, an area of overlap occurring between 24 and 56 Hz. Unpaired two-tailed t-test results showed that $\Delta F0$ differences between tone 2 (S2) and tone 3 (S1) were significant $[t(30)=4.21, p<0.001]$.

3. Discussion

The data in experiment 1 show that the two female speakers shared a region of overlap in the $F0$ range, and that tone 2 for the low-pitched speaker and tone 3 for the high-pitched speaker also overlap, two conditions essential to test for normalization effects. The hypothesis of this study was that listeners must normalize for speaker identity ($F0$ range) in order to identify these tones correctly. To test this hypothesis, perception of tone stimuli presented with high and low $F0$ precursors was examined, where tone stimuli formed a continuum from tone 2 to tone 3, varying in $\Delta F0$ and turning point characteristics.

II. EXPERIMENT 2: PERCEPTION OF TURNING POINT AND $\Delta F0$ IN ISOLATION

Although earlier studies using tones 2 and 3 for a single speaker have considered timing of the turning point and $\Delta F0$ to be perceptual cues for these two tones (e.g., Shen and Lin, 1991; Blicher *et al.*, 1990) these studies have not documented a systematic investigation of these parameters which addresses whether $\Delta F0$ and turning point covary, whether perception based on each of these parameters was equally categorical, or what combinations of turning point and $\Delta F0$ trigger shifts in identification from one tone to the other. A perception experiment was therefore devised to determine the relative importance of timing of the turning point and $\Delta F0$. The experiment tests perception of isolated synthetic stimuli in which timing of the turning point and $\Delta F0$ have been systematically manipulated. Subject responses should clarify how these acoustic dimensions are perceived, their relative importance, and any ambiguity created by the combination of parameters. In addition, results of this experiment were used to model more accurately the synthetic stimuli used in subsequent tests.

A. Method

1. Subjects

Six subjects from Mainland China, three males and three females between the ages of 19 and 40 years old, participated in experiment 2. All were recruited from the Cornell University community. None reported any hearing disorders. Subjects were paid for their participation.

2. Stimuli

The syllable [u] was chosen for synthesis because this syllable type was also used in the Shen *et al.* (1993) and Shen and Lin (1991) studies. Stimuli were created using the Delta speech synthesis program developed by Hertz (Charif *et al.*, 1992; Zsiga, 1994) and a Klatt synthesizer (1980) in the Cornell Phonetics Laboratory.[3] Formant frequency values

TABLE II. Formant frequency values (Hz) for synthesized stimuli.

	Onset	Offset
$F1$	345	304
$F2$	703	628
$F3$	2940	2940
$F4$	4320	4320
$F5$	4840	4840

for $F1-F5$ were averaged for the two speakers to create an ambiguous voice quality. Formant values for $F1$ and $F2$ included separate measurements for the onset and offset of the formants. Production data for $F3$ showed no substantial difference between onset and offset values for each speaker and thus $F3$ was held constant over the entire vowel. $F4$ and $F5$ were also held constant, based upon measurements from the steady-state portion of the vowel. The resulting composition of formant values is shown in Table II. Duration of the stimuli was constant at 400 ms. Amplitude of voicing began at 55 dB (Klatt parameter) and declined to 53 dB over the duration of the token.

Based on the turning point data in experiment 1, stimuli for the perception tests were designed to vary timing of the turning point along a continuum from 20 to 240 ms in 20-ms steps, for a total of 12 stimuli. These stimuli should trigger tone 2 responses when the turning point occurs close to the tone onset, and tone 3 responses when the turning point occurs late in the tone. In addition, $\Delta F0$ was varied from 10 to 70 Hz in steps of 5 Hz, generating 13 stimuli. Because tone 2 typically exhibits a shallower $\Delta F0$, it was expected that tones with a $\Delta F0$ equal to 10 Hz would produce more tone 2 responses than tones with a $\Delta F0$ of 70 Hz.

These two manipulations together allowed for testing of both timing of the turning point and $\Delta F0$, in an effort to understand how these acoustic parameters are used in the perception of tones 2 and 3. Figure 4 represents all combinations of these parameters which were included in experiment 2.

Based on the duration and $F0$ manipulations represented in Fig. 4, predictions can be made about which regions of the graph might be expected to trigger tone 2 and tone 3 responses. According to traditional phonetic descriptions, tone 2 is characterized by a short fall in $F0$ followed by a long rise, while tone 3 has a deeper, longer fall followed by a long rise. The shaded region in fig. 4 which corresponds to the tone 2 characterization contains stimuli with turning points from 20 ms to approximately 140 ms, along with lower $\Delta F0$ values. It might also be expected that a high $\Delta F0$ coupled with an early turning point (20 to 40 ms) would yield a tone 2 percept, since the $F0$ rise of the stimulus is predominant. On the other hand, listeners would be expected to label as tone 3 any stimulus containing a deeper $F0$ fall and a longer duration to turning point, stimuli marked in the lined region of Fig. 4.

3. Procedure

Experiment 2 used a forced-choice labeling paradigm in which subjects heard each [u] stimulus in isolation and were asked to choose from two lexical items. There were 468
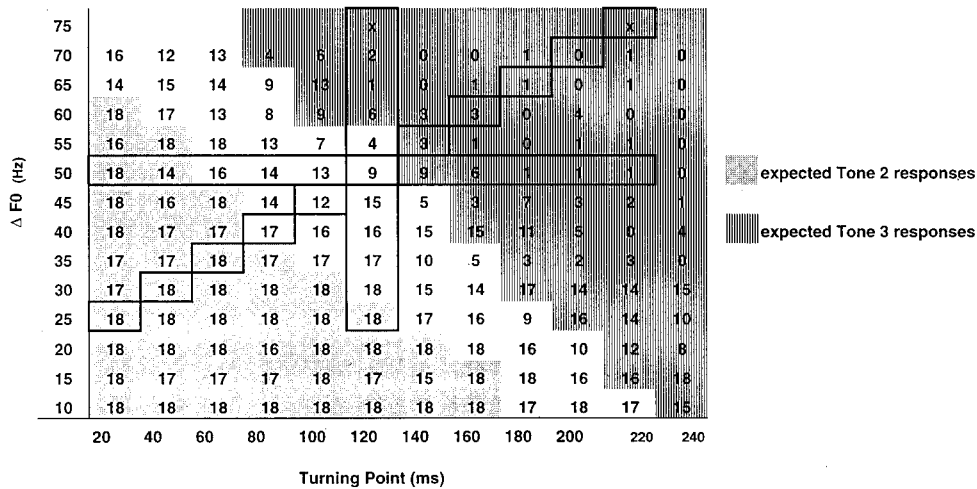
FIG. 4. Combinations of turning point and $\Delta F0$ manipulations for synthesized stimuli. Turning point manipulations are represented along the horizontal axis, $\Delta F0$ manipulations on the vertical axis. The shaded region corresponds to predicted tone 2 responses, and the lined region corresponds to predicted tone 3 responses. Numbers correspond to actual tone 2 responses in experiment 2 for isolated stimuli varying in timing of the turning point and $\Delta F0$. Eighteen responses were possible for each stimulus. These results determined which tone continua to use in experiments 3–5. These continua are enclosed in boxes: The diagonal boxes indicate stimuli varying along both turning point and $\Delta F0$, the horizontal boxes represent stimuli varying only in turning point, and the vertical row of boxes shows stimuli varying only in $\Delta F0$. An ''x'' denotes stimuli added after experiment 2.

tokens in total (12 turning point×13 $\Delta F0$ × 3 repetitions). Stimuli were low-pass filtered at 5.2 kHz and played out in randomized order on a 12-bit audio system using the BLISS software program (Mertus, 1989) on a Swan 386 PC. Due to the number of stimuli (156 different manipulations), stimuli were presented in three blocks, one set of stimuli per block, with an intertrial interval (ITI) of 2 s.

One to four subjects at a time participated in the test. They were instructed to respond to each item as quickly as possible by pressing the button corresponding to the Chinese character for ''not'' (tone 2) or ''dance'' (tone 3). To avoid misidentification, subjects were asked to pronounce the button labels prior to the experiment. A practice session consisting of 23 test items preceded the test. These practice items provided listeners with end points for each parameter manipulated, as well as stimuli in between. Instructions were given in English, since most of the subjects were undergraduate or graduate students at Cornell and highly proficient English speakers. However, for this and all subsequent tests, the few subjects who were not proficient in English were given instructions in Mandarin in addition to English. There were no differences in responses between the subjects instructed in Mandarin and those instructed in English. Subject responses were collected by computer using the BLISS software system. Responses for each stimulus were added across speakers.

B. Results

Figure 4 also gives the number of tone 2 responses for all stimuli, arranged according to values for timing of the turning point and $\Delta F0$. Figure 4 shows that, as expected, stimuli were clearly identified as tone 2 in the region where turning point and $\Delta F0$ values were low. Along the turning point dimension, unambiguous tone 2 responses spanned virtually the entire continuum, up to a $\Delta F0$ of 30 Hz. Thus tone

2 appears to tolerate substantial delays (up to 240 ms) when the initial $F0$ fall is 30 Hz or less. Although the total duration of the stimuli (400 ms) could have biased the listener toward tone 3 responses, this was clearly not the case.

Along the $\Delta F0$ dimension, decreases of up to 70 Hz were still identified as tone 2. According to Kratochvil (1971), who tested perception of synthetic tones with durations from 90 to 240 ms, a duration of between 50 and 100 ms is required for perception of isolated Mandarin tones. Weber ratios for duration have been reported for 100-ms signals as 0.026 by Ruhm et al. (1966), and for 400-ms signals as 0.12 by Stott (1935), corresponding to approximately 10–60 ms for the 400-ms stimuli in experiment 2. If the initial 20–40-ms portion of the tone is imperceptible (below threshold), subjects may hear only a rise, not the initial fall, since $\Delta F0$ is equivalent to 0 at the earliest turning points. $\Delta F0$ began to trigger tone 3 responses when the turning point reached approximately 80 ms into the tone. At this point, tone 3 responses increased as a function of both $\Delta F0$ and timing of the turning point. When the turning point occurred as late as 200 ms into the tone, however, relatively low $\Delta F0$ values (approximately 35 Hz or greater) elicited tone 3 responses. Thus the later the turning point, the easier it is for $\Delta F0$ to effect a change, though even late turning points are resilient to the effects of a $\Delta F0 < 35$ Hz. On the other hand, $\Delta F0$ appeared to trigger more tone 3 responses in stimuli with later turning points rather than early ones.

C. Discussion

The purpose of experiment 2 was to determine how the acoustic dimensions of timing of the turning point and $\Delta F0$ contribute to perception of Mandarin tones 2 and 3. As for which of the two acoustic dimensions might be more important, the data suggested that there is an interdependency between $\Delta F0$ and timing of the turning point. It ap-

pears that $\Delta F0$ was more relevant as turning point increased, while turning point was more relevant for a $\Delta F0$ of more than 30 Hz. Tone 2 perception seems to tolerate more variability overall, while tone 3 requires a late turning point *and* a large fall in $F0$. This may be because the later turning point enhances the perceptual salience of the initial fall.

While turning point and $\Delta F0$ appear to be interdependent, there were places where either dimension alone was sufficient to produce categorical functions. For example, stimuli with a constant turning point of 140 ms showed all tone 2 responses for a $\Delta F0$ of 20 Hz, moving to 50% responses for a $\Delta F0$ of 50 Hz, and all tone 3 responses for the largest $\Delta F0$. Along the turning point dimension, tone 2 responses moved categorically from 100% to 0% for the continuum of stimuli with a constant 50-Hz $\Delta F0$. While either of the two acoustic parameters were robust enough to trigger categorical identification functions, it is clear from both the production and perception data that timing of the turning point and $\Delta F0$ operate in tandem as perceptual cues to tones 2 and 3.

In summary, the results of this perception test showed how tone stimuli which vary in $\Delta F0$ and timing of the turning point are perceived. Subjects made categorical responses based on these two acoustic dimensions, such that identification functions were obtained for either $F0$, turning point, or both parameters. Experiments described in the following sections test normalization effects using tone 2 to tone 3 continua based on the acoustic parameters of $\Delta F0$ and turning point examined above.

III. EXPERIMENTS 3–5: PERCEPTION OF STIMULI IN PRECURSOR PHRASES

The following three experiments tested the hypothesis that listeners perceive tones in part by normalizing for speaker $F0$ range. Experiments 3–5 employed a design which compared how identical stimuli were identified when presented in two contexts differing in $F0$ range. To ensure that the effect was caused by normalization of different talker characteristics, the test used naturally spoken carrier phrases from different speakers as precursors. Normalization effects in this experiment would cause a shift in identification of stimuli as a function of which precursor phrase was heard, high or low $F0$.

As previously noted, earlier experiments on tones have provided evidence that extrinsic acoustic information may influence perception, although only Leather (1983) examined normalization effects due to perceived speaker identity. Experiments 3–5 of the present study expanded on Leather's work by testing two different Mandarin tones, tones 2 and 3, and sought to provide more robust evidence of normalization. This was done in several ways: by examining the direction of any shift in identification relative to the precursor, and by measuring shifts in identification over the entire function, rather than arbitrarily selected points in the middle.

The three experiments were conducted based on the perception data from experiment 2. Each experiment is distinguished according to the stimuli used: Experiment 3 employed a continuum of stimuli containing cues about both timing of the turning point and $\Delta F0$; stimuli for experiment

TABLE III. $F0$ and duration information for precursors used in experiments 3–5.

Speaker	Duration	Average $F0$	Peak	Valley
high $F0$	718 ms	226 Hz	272 Hz	192 Hz
low $F0$	722 ms	187 Hz	229 Hz	170 Hz

4 included a continuum of stimuli varying only $\Delta F0$, and experiment 5 used a continuum varying only timing of the turning point. The continua marked by boxes in Fig. 4 were used in these experiments. All three continua share a common midpoint which has a turning point of 120 ms, and a $\Delta F0$ of 50 Hz. This midpoint stimulus received 50% tone 2 responses for the corresponding identification functions resulting from experiment 2.

A. Experiment 3: Perception of stimuli varying in $\Delta F0$ and turning point

1. Method

a. Subjects. Eleven subjects, seven male and four female, aged between 19 and 40 years, participated in this experiment. All were native speakers of Mandarin Chinese, eight from Mainland China and three from Taiwan, with no known hearing disorders. Because there are many dialects spoken in Mainland China and Taiwan, the subject population in this study was restricted to those speaking only one of the Mandarin dialects according to Norman (1988, p. 191). This restriction provided a more homogeneous subject group, although not as strict as if they had been limited to Beijing Mandarin only. Examples of Chinese languages not represented by subjects included in the study were Shanghai, Cantonese, and Taiwanese.

b. Stimuli. Stimuli for this experiment were synthesized [u] syllables which formed a continuum from tone 2 to tone 3, varying in both timing of the turning point and $\Delta F0$. This continuum is the diagonal set of stimuli shown in Fig. 4. One additional step was created on the tone 3 end of the continuum to provide an equal number of stimuli on either end of the crossover stimulus. Timing of the turning point varied from 20 to 220 ms, in steps of 20 ms. $\Delta F0$ ranged from 25 to 75 Hz, in steps of 5 Hz.

Two natural precursor phrases spoken at a normal speaking rate were chosen from the production data discussed in experiment 1, one from each the high-pitched speaker (S1) and the low-pitched speaker (S2). Table III presents duration and $F0$ information for each precursor, high and low. The peak and valley $F0$ points represent boundaries of the $F0$ range for each speaker, showing a shared region of 192 to 229 Hz. The two phrases differed by 39 Hz in average $F0$, but were further distinguished by the range; the high precursor spanned 192–272 Hz, as compared to 170–229 Hz for the low precursor. In order to visualize how the stimuli were situated with respect to these $F0$ ranges, recall that the synthesized stimuli had a fixed onset and offset of 188 and 212 Hz, respectively, levels which were based upon production data. The $\Delta F0$ value decreased from 163 to 113 Hz in the continuum containing both $\Delta F0$ and turning point cues, and also in the $\Delta F0$ continuum.

Because they were naturally produced, the phrases differed in voice quality as well as $F0$ range. Although formant frequencies of the target stimuli were synthesized to be ambiguous relative to the precursors to reduce any speaker bias (see Table II), it could still be the case that one of the two precursors provided a better match with the target than the other.[4] Each phrase contained the segmental context *Zheige zi nian*— (''This word is—''), each had preceded a high-tone syllable (tone 1 or tone 4) in the production task,[5] and each matched in duration. Synthesized stimuli from the pre-test were appended to the precursors, leaving a 50-ms silence between the precursor and the test word.

As additional controls, the carrier phrase contained no instances of tones 2 or 3 or [u]. There were two advantages of limiting the phrases this way. One advantage was to eliminate the environment for tone sandhi effects, which particularly affect tones 2 and 3 (Chao, 1968). The other advantage is that subjects only heard one instance of the test tones, rather than possibly comparing precursor examples of the test tones with the stimuli.

2. Procedure and analysis

The experiment was conducted in the Cornell Phonetics Laboratory. Test items were presented to subjects by way of the PC-based software program BLISS, which randomized and played the stimuli via a D/A converter (12-bit resolution, 11-kHz sampling rate, low-pass filtered at 5.2 kHz). Eleven stimuli were preceded by each of the high and low precursor phrases, creating a total of 22 sentences. Subjects first heard 12 test items in a practice session. The test consisted of a total of 220 trials (22 sentences×10 repetitions), with an intertrial interval of 2250 ms. One to four subjects at a time listened to test items over headphones in separate booths. Subjects were instructed to respond by pressing one of two buttons, which were labeled using the Chinese characters for either ''not,'' or ''dance,'' corresponding to the tone 2 or tone 3 lexical item, respectively.

Responses were recorded and tabulated by computer. For each subject, the crossover points for stimuli in both the high and low precursor conditions were determined by using a probit statistical analysis (Finney, 1971). This statistical method takes into account the subject's responses over the entire stimulus continuum.

3. Results

Results of experiment 3 are summarized in figure and table form below. Figure 5 (top panel) shows the percentage of tone 2 responses for stimuli in the two presentation types (high and low precursor conditions), averaged across subjects. Table IV lists probit values for each subject as a function of the precursor conditions.

The probit values in Table IV represent category boundaries for the listeners who participated in experiment 3. As shown in Fig. 5, subjects perceived more tone 3 (low tone) responses when stimuli were preceded by a high precursor than when stimuli were preceded by a low precursor. The category boundary for the high precursor was earlier for eight of the eleven subjects, and averaged 5.60, as compared
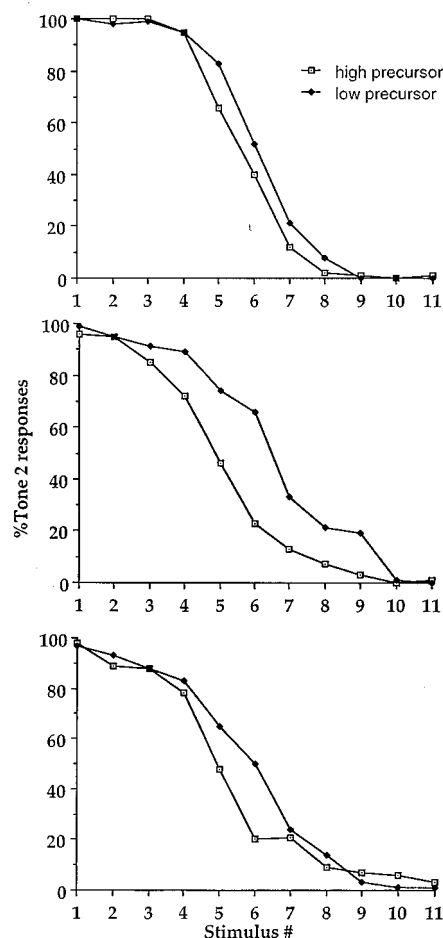
FIG. 5. Top panel: Experiment 3 turning point/$\Delta F0$ continuum identification functions for high and low precursor conditions, averaged across subjects. Stimulus 1 corresponds to predicted tone 2 responses. Middle panel: Experiment 4 $\Delta F0$ identification functions for high and low precursor conditions. Bottom panel: Experiment 5 turning point continuum identification functions for high and low precursor conditions.

to 5.99 for the low precursor. A paired two-tailed t test shows this difference between boundaries to be significant [$t(10) = -2.57$; $p < 0.03$]. Subjects thus appear to refer to the $F0$ range of the precursor in perception of the tones. Moreover, the normalization effect was robust enough to be obtained in a mixed block condition, in comparison to

TABLE IV. Experiment 3 probit values for turning point/$\Delta F0$ stimuli in high and low precursor conditions.

Subject	High precursor	Low precursor
1	5.58	5.21
2	4.64	4.51
3	5.71	6.2
4	6.34	7.11
5	5.88	5.81
6	4.36	5.65
7	5.75	6.54
8	5.3	5.44
9	6.18	6.99
10	6.39	6.71
11	5.44	5.63
mean	5.60	5.99

Subject	High precursor	Low precursor
1	4.47	6.18
2	4.74	6.08
3	6.02	7.76
4	4.29	4.09
5	3.33	4.75
6	4.49	4.56
7	3.32	6.28
8	4.79	6.57
9	5.06	7.24
10	4.10	5.59
mean	4.46	5.91

Leather (1983) in which stimuli were blocked by speaker.

The shift away from tone 2 responses in the high precursor condition demonstrates a contrast effect; the high $F0$ context caused a shift toward low tone (tone 3) responses. While this result is to be expected given the assumption that $F0$ height of the tone is interpreted relative to a speaker's $F0$ range, it differs from earlier findings by Fox and Qi (1990), who instead found primarily assimilatory shifts for paired-token identification tasks.

B. Experiment 4: Perception of stimuli varying in $\Delta F0$

1. Method

a. Subjects. Twenty-two native speakers of Mandarin Chinese participated in this experiment. Twelve of these were subsequently excluded from the results on the basis of criteria outlined below. The ten remaining subjects included five males and five females. One of the subjects was from Taiwan, and nine were from Mainland China. None reported any hearing disorders.

b. Stimuli. Test items were sentences composed of the two precursors used in experiment 3, followed by a test word taken from the $\Delta F0$ continuum in the pretest. This continuum varied only $\Delta F0$ in 11 steps of 5 Hz from 163 Hz. The timing of the turning point was fixed at 120 ms.

c. Procedure and analysis. The test procedure and analysis of data were identical to those used in experiment 3. However, this experiment seemed to be more difficult for subjects than experiment 3, judging both by number of missed trials and failure to achieve categorical identifications at continuum end points. Because of these two problems, it was decided that a subject's responses would be included in the results only if they met the following criteria: (1) they responded to more than 90% of the total trials for each continuum, and (2) they achieved at least an 80% correct response rate on continuum end points. Failure to meet these criteria led to the disqualification of 12 subjects.

2. Results

The averaged identification functions for the $\Delta F0$ continuum in the high and low precursor conditions are presented in Fig. 5 (middle panel) and the probit values are listed in Table V.

The data in Table V show that for the tone 2 to tone 3 continuum varying only in $\Delta F0$, there was an earlier shift to tone 3 responses in the high precursor condition for nine of the ten subjects; average crossover points were 4.46 as compared to 5.91 in the low precursor condition. This difference was significant $[t(9) = -4.69, p < 0.001]$. The shift was one of contrast—the high precursor prompted more low-tone responses and vice versa. This result supports the hypothesis that subjects refer to extrinsic $F0$ as a frame of reference for tone perception.

The magnitude of the shift in this experiment was much greater than that in experiment 3. These differences, computed as the difference between the low and high precursor crossover points, were shown to be significant in a two-tailed t test for independent means $[t(19) = -3.33, p < 0.003]$. These results indicate that listeners relied more on speaker $F0$ range to disambiguate the tones when the stimuli provided less intrinsic acoustic information about tone category. The implications of this finding are discussed in Sec. IV.

C. Experiment 5: Perception of stimuli varying in turning point

1. Method

Twenty native speakers of Mandarin Chinese participated in experiment 5. Eight subjects were disqualified according to the criteria outlined in Sec. III B. 1 c. The twelve remaining subjects included six males and six females. Four subjects were from Taiwan, and eight were from Mainland China. None reported any hearing disorders.

Again, stimuli used in this experiment were part of the set used in the isolation pretest, appended to the end of the natural precursors used in experiments 3 and 4. The continuum from tone 2 to tone 3 varied only timing of the turning point, from 20 to 220 ms into the tone. The decrease in $\Delta F0$ was constant at 50 Hz.

Test procedures and analysis of results were identical to those of experiments 3 and 4.

2. Results

Figure 5 (bottom panel) displays average percent tone 2 responses for the turning point continuum in the high and low precursor conditions. Table VI lists probit values for each subject in both the high and low precursor conditions. The average boundary in the high precursor condition as compared to the low was 4.78 versus 5.20. Only eight of the 12 subjects showed a shift in the direction predicted by the normalization hypothesis, and the difference between the probits in the two conditions was not statistically significant $[t(11) = -1.55, p > 0.15]$. These results suggest that stimuli varying only in a temporal cue did not induce a normalization effect for contexts that vary in an $F0$ dimension.

IV. GENERAL SUMMARY AND DISCUSSION

This study tested the hypothesis that listeners use acoustic information about the speaker in the perception of lexical tones. In particular, the study investigated whether listeners used extrinsic information about speaker $F0$ range in perception of intrinsic acoustic properties of Mandarin tones 2 and

TABLE VI. Experiment 5 probit values for turning point stimuli in high and low precursor conditions.

Subject	High precursor	Low precursor
1	5.83	5.15
2	5.54	6.12
3	5.25	4.73
4	4.55	4.86
5	4.21	5.39
6	4.0	5.08
7	4.56	4.44
8	3.24	5.63
9	4.13	5.12
10	3.61	4.01
11	6.08	6.34
12	6.4	5.49
mean	4.78	5.20

3. The hypothesis predicts that tone identification is affected by changes in perceived speaker identity. If speaker information is not relevant in tone perception, on the other hand, changes in perceived speaker identity should cause no significant shift in identification of tone categories.

To examine the hypothesis, a series of production and perception experiments were conducted. First, production analyses from experiment 1 located two speakers who overlapped in $F0$ range. The analysis revealed that within the area of $F0$ range overlap, a low tone for a high-pitched speaker and a high tone for a low-pitched speaker occurred at equivalent $F0$ heights. Experiment 2 demonstrated that while both $\Delta F0$ and turning point are used in production, either cue alone was sufficient to distinguish the two tones in perception.

The study then investigated whether changes in perceived speaker identity affected tone perception by presenting tone continua in precursor phrases from two different speakers. Results of experiments 3 and 4, which examined perception of both $F0$ and temporal properties of tones 2 and 3 in high-$F0$ and low-$F0$ precursor phrases, showed a small but significant shift in tone identification, in the direction expected if tone stimuli were perceived according to the $F0$ range of the precursor; that is, identical stimuli were perceived as high tones in the low $F0$ precursor phrase, but as low tones in the high $F0$ precursor phrase. These findings thus support the hypothesis that tone identification is influenced by changes in $F0$ range, demonstrating that this information is used as a frame of reference according to which ambiguous tones may be interpreted.

No significant shift was observed for the tone continuum in experiment 5, however, which varied only the temporal dimension of turning point. The stimuli in experiment 5 differed in only one aspect from the stimuli in experiments 3 and 4: they did not vary in $\Delta F0$. These results suggest that normalization is triggered only when both stimuli and precursors vary along the same acoustic dimension. Findings from Moore (1995) for tone stimuli identical to those used in the present study indicate that when precursors vary in a temporal dimension (speaking rate), listeners normalize for rate by shifting category boundaries for the temporal cue (turning point).

If the temporal dimension was not relevant for normalization for $F0$ range in the turning point stimuli, it is tempting to assume that temporal information may not have contributed to the normalization effect in experiment 3, where stimuli varied along both dimensions. However, the larger magnitude of the effect in experiment 4 as compared to experiment 3 contradicts this assumption. This difference in the magnitude of the effect for stimuli varying only in the $F0$ dimension as compared to stimuli varying in both the $F0$ *and* temporal dimensions supports the hypothesis that listeners utilize contextual information to a greater degree when intrinsic acoustic information for tone contrasts is degraded. Such differences between effects have been observed in rate normalization work on vowel perception by Gottfried *et al.* (1990), as well as rate effects in the perception of [b]-[w] continua in Shinn *et al.* (1985). Both of these studies show reductions in normalization effects as stimuli resemble natural speech more closely. Thus it is possible that when listeners are given accompanying temporal information for tone contrasts as in experiment 3, they do not refer to $F0$ range as much as when intrinsic tonal cues are restricted, as in experiment 4. Further work is needed to understand the relative contribution of temporal and $F0$ cues in contexts that also vary in both of these dimensions.

Although this investigation contributes additional data and addresses several inadequacies of Leather's (1983) study, findings of this study support the conclusions of Leather (1983). First of all, the present study observes $F0$ range normalization effects for tones which differ in $F0$ height; tone 2 is an upper register tone, compared to the lower register tone 3. Leather used two upper register tones whose contours are more dissimilar than tones 2 and 3. Second, this study shows normalization effects robust enough to be obtained in a mixed block condition; Leather's subjects were trained on one speaker's voice before hearing stimuli embedded in precursors for that particular speaker. Third, analysis methods for the present study compared crossover boundaries based on the entire identification function, so that reliable shifts were observable based on responses to all stimuli in each condition. The analysis of responses to only selected stimulus pairs rather than analysis of crossover boundaries may have led to the appearance of inconsistent results reported in Leather (1983) as well as Fox and Qi (1990). Fourth, while Leather did not report whether changes in perception were assimilatory or contrastive, or whether changes were consistent for all speakers, the present study provides conclusive evidence that shifts in identification were contrastive—in a direction opposite to the precursor $F0$—and that this shift was consistent across subjects in experiments 3 and 4.

The contrastive context effects shown here differ from those reported in Fox and Qi (1990). Their findings, for paired-token identification tasks, instead showed assimilatory shifts in all but one case. Their study focused on context effects from one preceding tone, rather than on speaker normalization, however. Fox and Qi further argued that assimilatory shifts are evidence for auditory, rather than phonetic, processing of the acoustic signal, based on experimental work by Fujisaki and Kawashima (1971), Pisoni (1975),

Shigeno and Fujisaki (1979), and Shigeno (1986). In these models of perception, assimilatory shifts occur for stimuli which do not undergo category-level perceptual identification, such as for continua whose end points do not represent different phonemes, or for nonspeech stimuli. For continua whose end points represent phonemic distinctions, or for complex tone continua, a categorical memory process is employed, generating contrastive shifts in identification. Shigeno (1991), however, provides evidence that both assimilatory and contrastive effects may occur within the process of phonetic judgment. From the standpoint of these two-stage perceptual processing models, results of the current study would suggest that higher-level phonetic processes are involved in speaker normalization for tones. Notwithstanding the different methods employed in Fox and Qi (1990) as compared to the present study, the opposite shifts in identification raise the question of whether contextual $F0$ information is processed differently depending upon whether it was used as a cue to tone identity, as in Fox and Qi, versus as a cue to speaker identity, as in the present study.

It could be argued that the present results arose from an auditory level of processing or from a simple response bias.[6] However, a comparison of the present results with those obtained from English listeners under the same experimental conditions argues against such explanations (Moore and Jongman, forthcoming). While the Mandarin listeners showed the greatest effect of speaker $F0$ range for stimuli varying only in $F0$, English listeners showed effects only when the foreign distinction was perceptually salient, namely when it was cued by both fundamental frequency and temporal information simultaneously. The fact that Mandarin and English listeners responded differently to our experimental conditions argues against a simple response bias. In addition, these findings suggest that normalization occurs for phonemic contrasts for native listeners, but that it is a function of auditory discriminability for non-native listeners.

Results of this study are consistent with those of Johnson (1990), who found that both intrinsic and extrinsic $F0$ contributes to vowel perception. As the results of experiments 3 and 4 from the present study demonstrate, extrinsic $F0$ significantly influences tone perception by serving as a cue to speaker identity, causing intrinsic $F0$ cues ($\Delta F0$) to be perceived relative to the extrinsic cues ($F0$ range). In other words, extrinsic $F0$ enabled listeners to construct a representation of $F0$ range, against which intrinsic acoustic characteristics of the tones were calibrated.

V. CONCLUSION

The results of this series of experiments suggest that perception of tones is a talker-contingent process. Evidence was provided to show that listeners use extrinsic $F0$ information in perception of lexical tones. Since no explicit tests were conducted to verify that the precursors were perceived as having been produced by different speakers, the present results cannot definitively demonstrate that listeners established a representation of speaker identity. However, the fact that the precursors in this study were natural, intact sentences produced by two different speakers, one with a high $F0$ and one with a low $F0$, suggests that intrinsic acoustic information is mediated through a representation of speaker identity, rather than contributing to tone identification independent of speaker information. These results suggest that the same normalization processes participate in perception of suprasegmentals as well as segments.

Speaker normalization has been assumed to occur as a response to acoustic variability which derives from vocal tract differences among speakers. This variability is exhibited when different speech sounds are acoustically very similar, as illustrated in this study, or when the same speech sound exhibits different acoustic characteristics. Additional research on normalization in perception in the latter instances of variability would further clarify the relationship between acoustic variability and normalization.

Other research on the effects of speaker variability on perception indicates that speech perception is more difficult, and not as accurate, in multiple-talker conditions as compared to single-talker conditions (Mullenix *et al.*, 1989; Sommers *et al.*, 1992), blocked conditions (Strange *et al.*, 1976; Assmann *et al.*, 1982) or when listeners have increased familiarity with the talkers' voices (Verbrugge *et al.*, 1976; Nygaard *et al.*, 1994). These studies suggest that there is a "cost" associated with the process of normalizing for speaker differences. While the costs of normalizing for contextual information may be expected for segments, which are perceived highly accurately given only intrinsic cues (Verbrugge *et al.*, 1976), it is not as straightforward in the case of suprasegmentals, where context is assumed to be more intimately connected with identification. In the case of Mandarin Chinese, contour differences between the tones also yield high identification rates in isolation (Howie, 1976). The more relevant case for establishing differing degrees of interdependence on context may be to examine normalization in perception of tones which contrast only in $F0$ height, such as the level tones in Cantonese (Fok, 1974). To the extent that tone perception uses identical perceptual processes as segments, the observation in the present study that normalization effects obtained in a mixed block condition suggests that speaker normalization is a robust process, even for Mandarin tones.

This study has illuminated the dual nature of tones as suprasegmentals in that both extrinsic and intrinsic acoustic information contribute to the description of a tone. Tones do not depend on absolute acoustic values to gain their identity. Rather, they contrast with other tones in the utterance as well as speaker $F0$ range to attain a relative identity. These assumptions are consistent with the results of this study showing that listeners use speaker $F0$ range in tone identification. Despite their intimate relationship with context, however, lexical tones also exist as independent phonological units, contrasting intrinsic acoustic characteristics such as turning point and $\Delta F0$ for Mandarin tones 2 and 3. Thus it is possible that in addition to contextual information specifying speaker $F0$ range, intrinsic $F0$ may also enable listeners to establish a representation of speaker identity. This hypothesis is consistent with findings by Slawson (1967) and Johnson (1990) for vowel perception, and Mullenix *et al.* (1989) for word recognition. The use of both extrinsic and intrinsic acoustic information in identifying speaker $F0$

range also avoids the ''bootstrap'' problem (Nearey, 1989), which confronts the issue of how listeners are able to establish a representation of speaker identity without precursor acoustic information.

ACKNOWLEDGMENTS

[1]Duration differences between the two tones may also be perceptually relevant (Blicher *et al.*, 1990), but will not be investigated in this study. Production data generally show that durations for both tones 2 and 3 are longer than for other tones and that tone 3 is longer than tone 2 (Dreher and Lee, 1966; Ting, 1971; Chuang *et al.*, 1972; Rumjancev, 1972; Lyovin, 1978; Nordenhake and Svantesson, 1983), perhaps because the nonprepausal form of tone 3 is shorter than in isolation.

[2]Other tokens of both tones 2 and 3 exhibited some degree of creak as well, although vowel onset and offset points were undisturbed. In these cases the creak was located in the middle of the vowel, and the expected formant structure returned before vowel offset.

[3]There has been much concern about whether female voices can be successfully synthesized given the current design of the Klatt synthesizer. Parameters now considered to improve the naturalness of synthesized female voices include breathiness, open quotient, and glottal waveform [see Klatt and Klatt (1990) for summary and experimental data]. Stimuli synthesized for the present experiment relied largely on manipulating traditional parameters of fundamental frequency and formant frequencies. In addition, a more breathy quality was modeled by setting a Delta parameter which filters the upper frequencies relative to the lower frequencies. Subjects reported hearing a female speaking, and were often surprised to learn the stimuli were not produced naturally.

[4]Although we have no direct indications as to whether our listeners perceived the precursor+target sequences as being produced by the same speaker, our instructions encouraged listeners to process the sequences as unitary. Specifically, listeners were told that they would hear sentences produced by two different speakers and that they were to listen to the entire sentence. After the experiments, the vast majority of listeners did not comment on any perceived mismatch between precursor and target, while only a few listeners remarked that one precursor made a more unitary sequence with the target than the other. Of course, in the latter case, it is unkown whether that impression was based on $F0$ or voice quality.

[5]This restriction to a high-tone context served as a control for tonal coarticulation cues which may have been present in carriers preceding tones 2 or 3. However, a study on coarticulation in Mandarin tones by Shen (1990) shows that there is no anticipatory effect on $F0$ height or direction from tones 2 or 3, and particularly no effect from those tones on a preceding tone 4. The similar $F0$ onset of tones 2 and 3 probably obviates coarticulation, since it would be in anticipation of the onset $F0$ height that anticipatory coarticulation would occur (Shen, 1990).

[6]We thank reviewer Rob Fox for raising these alternative interpretations.

Assmann, P., Nearey, T., and Hogan, J. (**1982**). ''Vowel identification: Orthographic, perceptual and acoustic aspects,'' J. Acoust. Soc. Am. **71**, 975–989.

Blicher, D. L., Diehl, R., and Cohen, L. B. (**1990**). ''Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: evidence of auditory enhancement,'' J. Phon. **18**, 37–49.

Fok, Chan Yuen-Yuen. (**1974**). ''A Perceptual Study of Tones in Cantonese,'' Occasional Papers and Monographs, Centre of Asian Studies (University of Hong Kong, Hong Kong), Vol. 18.

Chao, Y-R. (**1968**). *A Grammar of Spoken Chinese* (University of California, Berkeley).

Charif, R. A., Hertz, S. R., and Weber, T. J. (**1992**). *Delta System User's Guide* (Eloquent Technology, Ithaca, NY).

Chuang, C. K., Hiki, S., Sone, T., and Nimura, T. (**1972**). ''The acoustical features and perceptual cues of the four tones of Standard Colloquial Chinese,'' Proceedings of the Seventh International Congress on Acoustics, Budapest, p. 297–300.

Coster, D. C., and Kratochvil, P. (**1984**). ''Tone and stress discrimination in normal Beijing dialect speech,'' New Papers on Chinese Language Use (Canberra), p. 119–132.

Dreher, J., and Lee, P. C. (**1966**). ''Instrumental investigation of single and paired Mandarin tonemes,'' Res. Commun. **13**, Douglas Advanced Research Laboratories.

Finney, D. J. (**1971**). *Probit Analysis* (Cambridge U.P., Cambridge, England).

Fox, R., and Qi, Y. Y. (**1990**). ''Context effects in the perception of lexical tone,'' J. Chinese Ling. **18**, 261–283.

Fujisaki, H., and Kawashima, T. (**1971**). ''A model of the mechanisms for speech perception: Quantitative analysis of category effects in discrimination,'' *Annual Report of the Engineering Research Institute* (Faculty of Engineering, University of Tokyo), Vol. 30, pp. 59–68.

Gandour, J. (**1978**). ''The Perception of Tone,'' in *Tone: A Linguistic Survey*, edited by V. A. Fromkin (Academic, New York), pp. 41–76.

Gårding, E., Kratochvil, P., Svantesson, J.-O., and Zhang, J. (**1986**). ''Tone 4 and Tone 3 Discrimination in Modern Standard Chinese,'' Language Speech **29**, 281–293.

Gottfried, T. L., Miller, J. L., and Payton, P. E. (**1990**). ''Effect of speaking rate on the perception of vowels,'' Phonetica **47**, 155–172.

Howie, J. M. (**1976**). *Acoustical Studies of Mandarin Vowels and Tones* (Cambridge U.P., Cambridge, England).

Hsu, V. L. (**1990**). *A Reader in Post-Cultural Revolution Chinese Literature* (The Chinese University of Hong Kong, Hong Kong), pp. 344–381.

Johnson, K. (**1990**). ''The role of perceived speaker identity in $F0$ normalization of vowels,'' J. Acoust. Soc. Am. **88,** 642–654.

Johnson, K. (**1991**). ''Differential effects of speaker and vowel variability on fricative perception.'' Language Speech **34**, 265–279.

Jongman, A., and Miller, J. D. (**1991**). ''Method for the location of burst-onset spectra in the auditory-perceptual space: A study of place of articulation in voiceless stop consonants'', J. Acoust. Soc. Am. **89**, 867–873.

Kiriloff, C. (**1969**). ''On the auditory perception of tones in Mandarin,'' Phonetica **20**, 63–67.

Klatt, D. (**1980**). ''Software for a cascade/parallel formant synthesizer,'' J. Acoust. Soc. Am. **67,** 971–995.

Klatt, D., and Klatt, L. C. (**1990**). ''Analysis, synthesis and perception of voice quality variations among female and male talkers,'' J. Acoust. Soc. Am. **87,** 820–857.

Kratochvil, P. (**1971**). ''An experiment in the perception of Peking dialect,'' in *A Symposium on Chinese Grammar* (Scandanavian Institute of Asian Studies), pp. 7–31.

Ladefoged, P., and Broadbent, D. E. (**1957**). ''Information conveyed by vowels,'' J. Acoust. Soc. Am. **29,** 98–104.

Leather, J. (**1983**). ''Speaker normalization in perception of lexical tone,'' J. Phonetics **11**, 373–382.

Li, C., and Thompson, S. (**1981**). *Mandarin Chinese: A Functional Reference Grammar* (University of California, Berkeley).

Li, C. N., and Thompson, S. (**1978**). ''The acquisition of tone in Mandarin-speaking children,'' in *Tone: A Linguistic Survey*, edited by V. A. Fromkin (Academic, New York).

Lin, M. C. (**1988**). ''Putong hua sheng diao de sheng xue texing he zhi jue zhengzhao, [Standard Mandarin tone characteristics and percepts],'' Zhongguo Yuyan **3**, 182–193.

Lin, T., and Wang, W. Y.-S. (**1985**). ''Shengdiao ganzhi wenti [tone perception],'' Zhongguo Yuyan Xuebao **2**, 59–69.

Lyovin, A. (**1978**). ''Review of Tone and Intonation in Modern Chinese by M. K. Rumjancev,'' J. Chinese Ling. **6**, 120–168.

Mann, V. A., and Repp, B. H. (**1980**). ''Influence of vocalic context on the perception of the [ʃ]-[s] distinction,'' Percept. Psychophys. **23**, 213–228.

Mertus, J. (**1989**). *BLISS Manual* (Brown University, Providence, RI).

Moore, C. B. (**1993**). ''Some observations on tones and stress in Mandarin

Chinese,'' Working Papers of the Cornell Phonetics Laboratory **8**, 82–117.

Moore, C. B. (**1995**). ''Speaker and rate normalization in the perception of lexical tone by Mandarin and English listeners,'' Ph.D. dissertation, Cornell University.

Moore, C. B., and Jongman, A. (forthcoming). ''Cross-language effects in the perception of Mandarin tone.''

Mullennix, J. W., Pisoni, D. B., and Martin, C. S. (**1989**). ''Some effects of talker variability on spoken word recognition,'' J. Acoust. Soc. Am. **85,** 365–378.

Nearey, T. M. (**1989**). ''Static, dynamic, and relational properties in vowel perception,'' J. Acoust. Soc. Am. **85,** 2088–2113.

Nordenhake, M., and Svantesson, J.-O. (**1983**). ''Duration of Standard Chinese word tones in different sentence environments,'' Working Papers **25** (Lund, Sweden), 105–111.

Norman, J. (**1988**). *Chinese* (Cambridge U.P., Cambridge, England).

Nygaard, L. C., Sommers, M. S., and Pisoni, D. B. (**1994**). ''Speech perception as a talker-contingent process,'' Psychol. Sci. **5**, 42–46.

Peterson, G. E., and Barney, H. L. (**1952**). ''Control methods used in a study of the vowels,'' J. Acoust. Soc. Am. **24**, 175–184.

Pisoni, D. B. (**1975**). ''Auditory short-term memory and vowel perception,'' Mem. Cogn. **3**, 7–18.

Ruhm, H. B., Mencke, E. O., Milburn, B., Cooper, Jr., W. A., and Rose, D. E. (**1966**). ''Differential sensitivity to duration of acoustic stimuli,'' J. Speech Hear. Res. **9**, 371–384.

Rumjancev, M. K. (**1972**). ''Ton i Intonacija v Sovremennon Kitajskom Jazyke [Tone and Intonation in Modern Chinese] (Izdatel'stvo Moskovskogo Universiteta, Moscow),'' reviewed by A. V. Lyovin (1978). J. Chinese Ling. **6**, 120–168.

Shen, X. (**1990**). ''Tonal coarticulation in Mandarin,'' J. Phonetics **18**, 281–285.

Shen, X., and Lin, M. (**1991**). ''A perceptual study of Mandarin Tones 2 and 3,'' Language Speech **34**, 145–156.

Shen, X., Lin, M., and Yan, J. (**1993**). ''*F*0 turning point as an *F*0 cue to tonal contrast: A case study of Mandarin tones 2 and 3,'' J. Acoust. Soc. Am. **93,** 2241–2243.

Shigeno, S. (**1986**). ''The auditory tau and kappa effects for speech and nonspeech stimuli,'' Percept. Psychophys. **40,** 9–19.

Shigeno, S. (**1991**). ''Assimilation and contrast in the phonetic perception of vowels,'' J. Acoust. Soc. Am. **90,** 103–111.

Shigeno, S., and Fujisaki, H. (**1979**). ''Effect of a preceding anchor upon the categorical judgment of speech and nonspeech stimuli,'' Jpn. Psychol. Res. **21**, 165–173.

Shinn, P. C., Blumstein, S. E., and Jongman, A. (**1985**). ''Limitations of context conditioned effects in the perception of [b] and [w],'' Percept. Psychophys. **38**, 397–407.

Slawson, A. W. (**1967**). ''Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency,'' J. Acoust. Soc. Am. **43,** 87–101.

Sommers, M. S., Nygaard, L. C., and Pisoni, D. B. (**1992**). ''Stimulus variability and the perception of spoken words: Effects of variations in speaking rate and overall amplitude,'' in *ICSLP 92 Proceedings: 1992 International Conference on Spoken Language Processing*, edited by J. J. Ohala, T M. Nearey, B. L. Derwing, M. M. Hodge, and G. E. Wiebe (Priority Printing, Edmonton, Canada), Vol. 1, pp. 217–220.

Stott, L. H. (**1935**). ''Time-order errors in the discrimination of short tonal durations,'' J. Exp. Psychol. **18,** 741–766.

Strange, W., Verbrugge, R., Shankweiler, D., and Edman, T. (**1976**). ''Consonant environment specifies vowel identity,'' J. Acoust. Soc. Am. **60,** 213–224.

Ting, A. C. (**1971**). ''Mandarin tones in selected sentence environments: An acoustic study,'' Ph.D. dissertation, University of Wisconsin.

Verbrugge, R. R., Strange, W., Shankweiler, D. P., and Edman, T. R. (**1976**). ''What information enables a listener to map a talker's vowel space?,'' J. Acoust. Soc. Am. **60,** 198–212.

Whalen, D. H. (**1981**). ''Effects of vocalic formant transitions and vowel quality on the English [s]-[ʃ] boundary,'' J. Acoust. Soc. Am. **69,** 275–282.

Zsiga, E. (**1994**). *Syllt: The Delta Syllable Tool* (Eloquent Technology, Ithaca, NY).

Acoustics, perception, and production of *legato* articulation on a computer-controlled grand piano

Bruno H. Repp[a)]

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511-5595

In an attempt to replicate and extend previous results obtained on a digital piano [B. H. Repp, J. Acoust. Soc. Am. **97**, 3862–3874 (1995)], the present study analyzed piano tone decay characteristics, musically trained listeners' interactive adjustments of key overlap times (KOTs) in tone sequences, and samples of pianists' *legato* playing in scales and *arpeggi* on a computer-controlled Yamaha Disklavier. On the whole, the results resembled the earlier findings: In both perception and production, KOTs tended to be longer for high rather than for low tones and for relatively consonant rather than for dissonant successive tones. KOTs also increased as tempo decreased in production, but there was no corresponding effect in perception (where a smaller range of tempi was used). Even though the decay times of the natural piano tones were about twice as long as those of the digital piano used earlier, the average KOTs were not shorter; on the contrary, they were longer in the perception task, while there was little difference in production. Perception of optimal *legato* does not seem to rest on an invariant criterion of acoustic tone overlap, and pianists do not seem to adjust their KOTs substantially when playing on different instruments. However, there were large individual differences in KOTs. © *1997 Acoustical Society of America.* [S0001-4966(97)00809-6]

PACS numbers: 43.75.St, 43.75.Mn, 43.75.Cd, 43.66.Mk [WJS]

INTRODUCTION

Differences in articulation are an important ingredient of technically competent and expressive performance on many instruments. The term ''articulation'' refers to the manner in which tones are initiated, terminated, and connected to following tones. Keyboard players, unlike string and wind instrument players, do not have much control over the onset characteristics of tones, so that articulation on the piano is mainly a matter of connecting or separating consecutive tones. Connected tones are produced by releasing a key at or slightly after the moment at which another key is depressed.[1] This articulation is called *legato*, and the amount of overlap of the two successive key depressions will be referred to as *key overlap time* (KOT). A negative KOT indicates that the first key was released before the second key was depressed. If this gap is long enough, it will cause the two tones to be perceived as unconnected or *nonlegato*, and if the first tone is in addition rather short, its articulation is referred to as *staccato*.

The present study complements an earlier study of *legato* articulation on a computer-controlled digital piano (Repp, 1995). That instrument (a Roland RD-250s) produced synthetic piano tones that did not have all the variability of natural piano tones. The keyboard was weighted but did not have the exact feel of a real piano, and the sound output was heard over earphones. For these reasons, the generalizability of the findings to real pianos was not assured. It was the purpose of the present study to repeat the earlier measurements and experiments on a computer-controlled grand piano (Yamaha Disklavier). Relevant earlier literature, including

the pioneering study of Kuwano *et al.* (1994), is reviewed in Repp (1995).

Repp's study had three parts. First, an acoustic analysis of the tones produced by the digital piano was conducted, in order to describe the sonic environment within which the following perception and production experiments took place. Of main interest were the decay characteristics of the tones, especially the decay following key release, called the *postrelease decay*. In a real piano, this decay includes not only the decay of the damped string vibrations but also the decay of soundboard vibrations and of acoustic reverberation in the room. There is little information about postrelease decay times in the piano acoustics literature. Repp's measurements indicated that these decay times can be up to several hundreds of milliseconds long and that low tones decay more slowly than high tones (as they do prior to key release). However, it is not clear how representative these decay times are of real piano tones, and since they may be instrument- and room-specific, they had to be measured in the present experimental setting as well.

The second part of Repp's study was a perceptual experiment in which musically trained participants (including pianists) listened to repeatedly ascending and descending sequences of tones played under computer control on the digital piano and adjusted their KOTs interactively until the sequences sounded optimally, minimally, or maximally *legato*. The listeners' adjustments in the three conditions provided information about the range of KOTs acceptable as *legato*. The tone sequences varied in tempo, register (pitch height), and pitch step size. Since short piano tones, because of their greater intensity at the time of key release, have longer postrelease decay times than long tones, and since low tones have longer decay times than high tones, it was predicted

---

that adjusted KOTs would be shorter at a faster tempo and in a lower register, to compensate for the greater acoustic overlap due to the longer postrelease decay. It was also predicted that tones separated by 3 semitones (st) would yield longer KOTs than tones separated by 1 st, due to their greater consonance. Support was found for all three predictions, though there were some interactions that will be discussed in greater detail below.

In the third part of the earlier study, pianists were recorded playing on the digital piano, with optimal *legato* articulation, ascending and descending tone sequences similar to the ones used in the perceptual experiment. Their KOTs were determined from the Musical Instrument Digital Interface (MIDI) recordings. KOTs increased in absolute magnitude as tempo decreased, an effect that is consistent with the perceptual results, although it may well have a kinematic rather than perceptual origin. There was no effect of register, however, suggesting that the pianists did not adjust their KOTs to the different postrelease decay times of high and low tones. KOTs did tend to increase with the relative consonance (pitch step size) of the tones. The data also revealed large individual differences in average KOT (a range from 27 to 145 ms), a tendency of some pianists to play more *legato* with the right hand than with the left (in the same register), and, most interestingly, a highly systematic pattern of KOTs as a function of position in the sequence: KOTs were longest at the beginning of a unidirectional sequence and decreased as the sequence progressed, increasing again after a reversal of direction. This suggested that, even within the *legato* range, KOTs vary systematically with the metrical or melodic structure of the sequence.

The present study was divided into three analogous parts. Its design was similar to, but not identical with, that of the earlier study. The goal of the initial acoustic analyses was to determine the pre- and postrelease decay characteristics of the natural piano tones used in the subsequent experiments. In the perceptual part, listeners controlled KOTs interactively on a Yamaha Disklavier in order to find the optimal, minimal, or maximal *legato* for tone sequences varying in tempo, register, and pitch step size. In the production part, pianists played the tone sequences used in the perceptual experiment, and their KOTs were recorded and analyzed to determine whether adjustment to different amounts of acoustic overlap occurred, as well as to observe other systematic patterns and individual differences. Finally, in a task not included in the earlier study, the pianists were asked to perform a short musical composition involving *legato* articulation, in order to check whether some of the findings extended to the playing of real music.

I. ACOUSTIC ANALYSIS

A. Method

The acoustic measurements to be reported were conducted on the tones of a Yamaha Disklavier Mark II grand piano, the instrument used in the production experiment.[2] It was less than 1 year old at the time of recording and stood in the center of a large classroom, with its lid raised to the highest position. The recording microphone (Shure BG 5.0 cardioid) was pointed at the piano from the right side at a

height of about 5 ft, about 3 ft from the curved rim of the frame. Some additional recordings were made from a position in front of the piano, approximately where a pianist's head would be. All recordings were made on digital audio tape.

The tones were produced under the control of a simple program written in MAX (a graphic programming language for MIDI applications), running on a Macintosh computer. The solenoid-driven piano keys were played in ascending order, with nominal intertone (key release to key depression) intervals of 1 s. This allowed each damped tone to die away completely before the next tone started. Tones with a nominal (key depression to key release) duration of 250 ms were played at five MIDI velocities (relative intensities): 20, 40, 60, 80, and 100. At the MIDI velocity of 100, they were also played with two longer nominal durations: 500 and 1000 ms.[3] (Nominal duration was meaningless for tones above F6 because of the absence of dampers.) As in Repp (1995), the analysis was restricted to the five octaves between C2 (65 Hz) and C7 (2093 Hz), but whereas only every third tone in the chromatic scale had been analyzed in the previous study, all tones were measured here.

To conduct these measurements, the digital recordings were resampled by a Macintosh Quadra 660 AV computer at 22.055 kHz/s. SIGNALYZE software was used to compute the root-mean-square amplitude envelope of each tone with a running rectangular integration window. The width of the window was set to 40 ms for tones in the lowest two octaves, but from C4 (262 Hz) to C7 (2093 Hz) it was progressively reduced in millisecond steps to encompass approximately ten fundamental periods. The window duration for the highest tone thus was 5 ms. The reason for this procedure was that, as pitch (and with it, sustained decay rates) increased, the energy peak of the tones became narrower, so that a fixed integration window would have led to an artifactual decrease in peak sound level with pitch. (A fixed 30-ms window had been used by Repp, 1995.)

The amplitude at the envelope peak was determined and recorded. Decay during the sustained portion (prerelease decay) was measured in the 1000-ms tones by determining the envelope amplitude every 100 ms from the peak. Postrelease decay was measured by determining the amplitude at the time of key release (i.e., with the integration window centered on that point in time) and every 50 ms from then on, until the amplitude was near the noise floor.[4] All amplitude values were subsequently converted to sound levels in dB, with an arbitrary reference.

B. Results and discussion

Figure 1 shows the basic relation between MIDI velocity and peak sound level, averaged across all pitches. This relationship was virtually constant across the whole pitch range and was best approximated by a quadratic (not a logarithmic) function. For comparison, the very similar function obtained for the Roland RD-250s digital piano used in Repp (1995) is shown.[5] At low MIDI velocities, a 1-dB change equalled about 2.5 velocity units; at high MIDI velocities, about 5 velocity units. The dynamic range represented is about 23 dB. The full range is likely to be a bit wider, but not by much
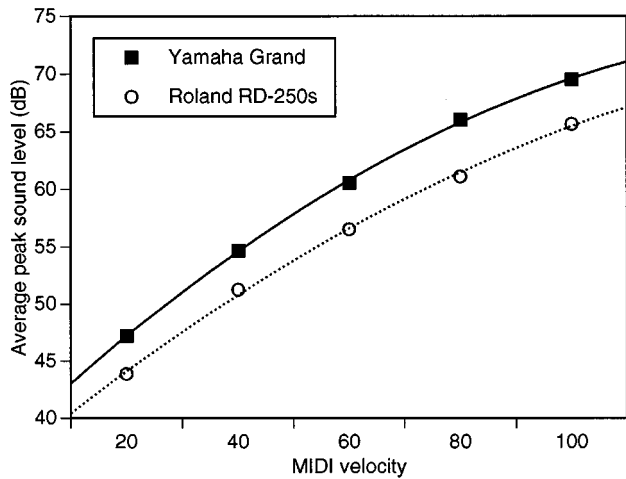
FIG. 1. The increase in average peak sound level with MIDI velocity on the Yamaha Disklavier grand piano and on the Roland RD-250s digital piano (''Piano 1'' sound). The curves are quadratic functions fitted to the data points.

FIG. 3. Prerelease decay contours of tones with a MIDI velocity of 100 and a nominal duration of 1000 ms: Sound levels at the peak (heavy line) and every 100 ms after the peak (solid and dotted lines alternating), until the nominal offset.

because the Yamaha Disklavier does not respond monotonically towards the extremes of the MIDI velocity range (0–127).

Figure 2 shows peak sound levels for all tones as a function of pitch and MIDI velocity. (Two keys in the highest octave failed to produce a sound at the lowest MIDI velocity.) These data suggest that peak intensity was slightly higher in the middle register but did not change very systematically across the pitch range.[6] However, there was large unsystematic variation in peak sound level from one tone to the next, as was observed in earlier measurements on two acoustic instruments (Repp, 1993). The origin of this variation had remained unclear in the earlier study, where only a single microphone position had been used. In the present study, an additional set of tones with a MIDI velocity of 60 was recorded from a position in front of the piano. The dotted line in Fig. 2 represents the peak sound levels of these tones. Clearly, the pattern of the unsystematic variation was

largely specific to the location of the microphone. The correlation between the two functions for a MIDI velocity of 60 is only 0.28 ($p < 0.05$), whereas that between two functions for the same microphone location (MIDI velocities of 60 and 80) is 0.97. Evidently, then, the irregularities do not originate in the piano mechanism but rather are due to sound radiation in the room (see Benade, 1976).

Next we consider the sustained (or prerelease) decay of the piano tones. Figure 3 shows the peak sound levels at a MIDI velocity of 100 (heavy line), as well as the sound levels of 1000-ms tones measured every 100 ms after the peak, up to the nominal offset (key release). The figure confirms the well-known fact (Martin, 1947) that higher piano tones decay much faster than lower tones. It also shows that the decay pattern of tones above C5 is very different from that of tones below C5: While lower tones seem to decay at a fairly steady rate (over the first second, at least), higher tones rapidly lose energy (about 15 dB) over the first 200 ms and then decay in a highly irregular and nonmonotonic manner.

Figure 4 portrays the postrelease (damped) decay characteristics for tones with a nominal duration of 250 ms and a MIDI velocity of 100. The heavy line on top shows again the peak levels, and the heavy line immediately below it shows the sound levels at the nominal offset (key release). The remaining functions represent the sound levels every 50 ms after the offset, down to the ambient noise floor. It is evident that the postrelease decay rates do not vary as much with pitch as do the prerelease decay rates, but there are again irregularities at the higher pitches. Tones above F6, which have no dampers, decay very little and very irregularly (cf. also Fig. 3). Similar graphs were obtained for 500- and 1000-ms tones but are not shown here to conserve space.

Estimates of pre- and postrelease decay rates were obtained by fitting linear functions to the decay measurements of each tone, even if the decay was very irregular. The absolute values of the slopes of the linear functions thus provide estimates of average decay rates, over about 900 ms
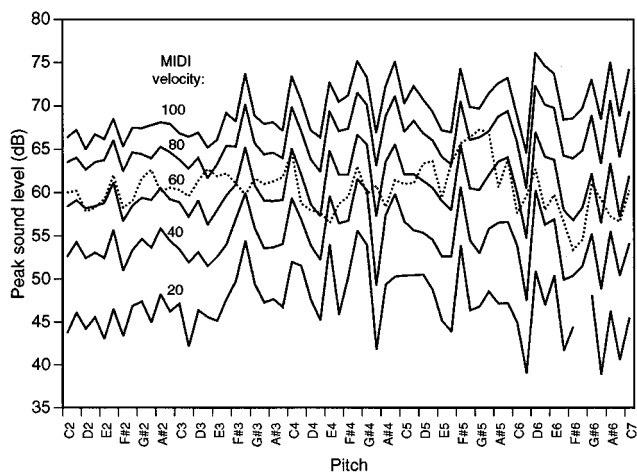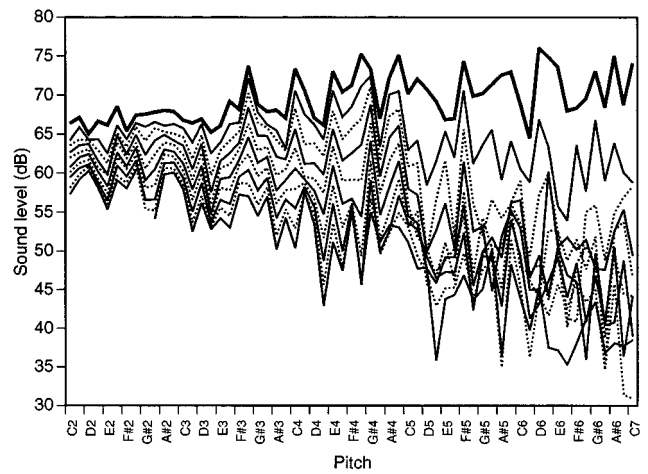

FIG. 2. Peak sound level as a function of pitch and MIDI velocity. The dotted line represents tones with a MIDI velocity of 60, recorded with the microphone in a different position. Only every other pitch label is shown on the abscissa.
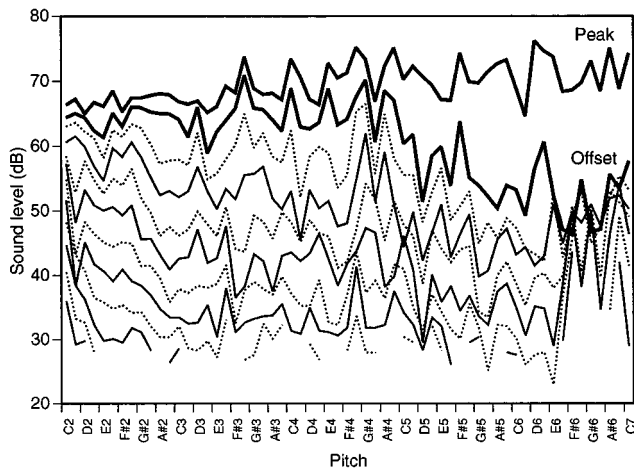
FIG. 4. Postrelease decay contours of tones with a MIDI velocity of 100 and a nominal duration of 250 ms: Sound levels at the peak and at the nominal offset (heavy lines), and every 50 ms from the offset (dotted and solid lines alternating), down to the ambient noise floor.

(sampled every 100 ms) in the case of sustained tones, and over about 500 ms (sampled every 50 ms) in the case of damped tones. Figure 5 summarizes these decay rates, which are expressed in dB per centisecond (cs).

The prerelease decay rates (for 1000-ms tones) increased roughly as a linear function of pitch, from about 1 dB/cs to nearly 3 dB/cs, though there was considerable variability from tone to tone, as already observed in earlier studies (Martin, 1947; Hundley *et al.*, 1978). Similar data for a Baldwin grand piano reported by Hundley *et al.* (1978) showed an increase in decay rates from about 1 dB/cs (C2) to about 10 dB/cs (C7). These authors may have been measuring initial decay rates only or used a piano that was tuned with exceptional accuracy. As Fig. 3 indicates, the decay functions of the present tones were quite linear up to C5, but tones higher than C5 showed a much faster (also apparently linear) decay over the first few hundreds of milliseconds,

followed by a slower and usually highly irregular decay afterwards. Hundley *et al.* (1978) have shown that the two decay rates result from interference among the multiple strings, due to slight differences in tuning that cause a loss of phase synchrony over time, which results in reduced energy transmission to the soundboard and in beats that modulate the decay (cf. Fig. 3). The lower tones showed a single decay rate here because it usually takes more than one second for their strings to get out of phase.[7] If measured over a longer time, they typically show two decay rates also (Martin, 1947).

Naturally, the postrelease (damped) decay rates were much faster than the prerelease (sustained) decay rates, but there was again remarkable variability from tone to tone. Moreover, while 500- and 1000-ms tones seemed to decay at about the same rates, 250-ms tones decayed at a faster rate, for reasons that are not immediately clear. For the longer tones, there was a tendency for postrelease decay rates to decrease as pitch increased, although the lowest tones in the range also exhibited slower decay rates. Presumably, the postrelease decay rates reflect sound reverberation in the acoustic space; the decay of the string and sound board vibrations giving rise to the sound is probably a good deal faster.

The prerelease decay rates of the digital piano tones measured by Repp (1995: Fig. 1) were similar to those of the present Yamaha Disklavier tones, although there was evidence of two decay rates only from C6 on (as compared to C5 in the present data). However, the postrelease decay rates of the digital piano tones (Repp, 1995: Fig. 3) were more than twice as fast as those of the Yamaha tones: They increased from about 11 to 18 dB/cs during the lowest (second) octave, and then increased more slowly to about 22 dB/cs in the sixth octave. Evidently, the digital piano tones simulated a less reverberant environment. This is relevant to a comparison of the earlier results with the results of the following experiments.

II. PERCEPTION EXPERIMENT

The purpose of this experiment was to determine whether the earlier perceptual results obtained on a digital piano can be generalized to a real acoustic instrument. As in Repp (1995), the task required listeners to adjust key overlap times (KOTs) interactively, so as to achieve a specified degree of *legato* in repeatedly ascending and descending tone sequences. However, the design differed from that of the earlier study in three ways: (1) To vary tempo, the earlier experiment had used tone interonset intervals (IOIs) of 260, 520, 779, and 1039 ms.[8] On the Yamaha Disklavier, the shortest IOIs presented a problem because at the melodic turning points of sequences, where the same tone occurred twice with only one other tone intervening, there was not enough time for the key to be released and depressed again when a long KOT was selected. To avoid missing tones in the sequences, the fastest tempo was omitted and only the three longer IOIs were used. (2) To make up for this reduction in the number of conditions, three rather than two different pitch step sizes were used. (3) Finally, in order to eliminate an earlier confounding of step size with the aver-
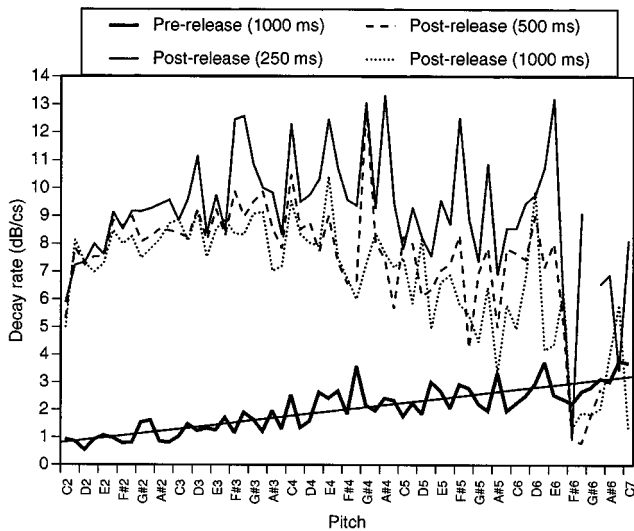

FIG. 5. Average prerelease decay rates for 1000-ms tones and postrelease decay rates for tones of 250-, 500-, and 1000-ms duration. A linear function has been fitted to the prerelease decay rates.

age pitch of the tones in a sequence, sequences differing in step size were centered on the same pitch. Three registers (octaves) were used, as in the earlier study.

The predictions were the following: (1) The KOT chosen to represent the optimal *legato* should increase with register because damped high tones decay faster and hence show less acoustic overlap than damped low tones. (2) The KOT should increase with IOI because long tones are softer than short tones at the time of key release and therefore decay sooner and overlap less. However, this prediction was confirmed only at the short IOIs (260 versus 520 ms) in the earlier study, and since the shortest IOI was omitted here it was not clear whether an effect of IOI should be expected. (3) The KOT should increase with step size because the successive tones become increasingly consonant at the pitch steps chosen (1, 2, and 3 st, corresponding to intervals of a minor second, major second, and minor third). (4) All these effects should be enhanced when KOTs are chosen to represent maximal *legato*. (5) According to the earlier results, the (generally negative) KOTs chosen to represent minimal *legato* should primarily show an effect of IOI: The longer the IOI, the longer the nominal gap that can be tolerated between successive tones. (6) All these effects will have to be fairly robust to emerge from the considerable variability seen previously in listeners' perceptual judgments. This variability was expected to be increased further by the unsystematic acoustic variation inherent in natural piano tones and the unpredictable effects of room acoustics.

A. Method

1. Listeners

Repp (1995) compared pianists with nonpianists and did not find any systematic differences in their judgments. Therefore the present listeners were only required to have some musical training and thus presumably a concept of *legato*. Fifteen individuals participated, 13 of whom were of college age, recruited by advertisement, and paid for their services; the other two were the author and his research assistant. A wide range of musical experience was represented, ranging from informal music making to more than 10 yr of training on an instrument. Five listeners were pianists. The data of four of the less experienced listeners were excluded because of extreme variability or anomalous responses (large negative KOTs for optimal *legato* in one case, large positive KOTs for minimal *legato* in another). This left data from 11 listeners to be analyzed.

2. Design

The stimuli were continuously ascending and descending 5-tone sequences. There were 27 conditions resulting from the combination of three IOIs, three registers, and three step sizes. The IOIs were 520, 779, and 1039 ms. The registers were the second, fourth, and sixth octave. The step sizes were 1, 2, and 3 st. To keep the average pitch constant within each register, all sequences were centered on G. The 1-st sequence thus consisted of the pitch classes F-F#-G-G#-A, the 2-st sequence of D#-F-G-A-B, and the 3-st sequence of C#-E-G-A#-C#.

3. Procedure

A Yamaha Disklavier Mark II baby grand piano, standing in the same place as the grand piano that eventually replaced it, was controlled by a MAX program running on a Macintosh computer.[9] The listener sat in front of the computer, which stood on a table about 6 feet to the right of the piano, and started and stopped sequences by clicking on "buttons" displayed on the monitor. Also displayed was a "slider" with a "handle" that could be dragged or clicked with the mouse to vary KOT. The slider actually controlled the duration of the tones within a fixed nominal range (250 to 1500 ms). The point of zero KOT on the slider thus varied with the IOI. The slider was unlabeled and was automatically reset to the left-most position (250 ms) at the beginning of each trial, so that the tones in each sequence initially sounded short and unconnected. The listener's task was to start a trial, move the slider first to the right and then back and forth until the tone sequence satisfied the specified perceptual criterion, then to stop the sequence and start the next trial. Each sequence started on its lowest pitch and continued ascending and descending until it was stopped. The MAX program presented the 27 sequences in random order and saved the listener's final slider settings.

After some initial practice, each listener completed four blocks of 27 trials each. In block 1 the task was to find the best *legato* in which the tones seemed "optimally connected." In block 2 it was to find the minimal *legato* in which the tones were "barely connected" but not separated by audible gaps. In block 3 it was to find the maximal *legato* that was still acceptable, without annoying overlap of tones being heard. In block 4 the listeners once again tried to determine the optimal *legato*. Blocks were separated by short breaks. Listeners typically took about 30 s for an adjustment, but there were considerable individual differences in pace. Head movements were not constrained, and listeners were encouraged to turn their head towards the piano while making their KOT adjustments.

B. Results and discussion

The average KOTs chosen by the 11 listeners are shown in Fig. 6, which may be compared to the results on the digital piano shown in Fig. 4 of Repp (1995). The error bars represent ±1 standard error. The results of blocks 1 and 4, which had been averaged in the earlier study, are displayed separately here (as "best 1" and "best 2," respectively) because a change in criterion was evident: Listeners chose longer optimal KOTs (by 54 ms, on the average) at the end of the experiment than at the beginning. The difference was significant in a four-way (block, register, step size, IOI) repeated-measures analysis of variance (ANOVA) on the best *legato* data [$F(1,10) = 10.67$, $p < 0.009$]. (A trend in the opposite direction had been found in the earlier experiment.) There was large between-listener variability, comparable to that observed in the earlier study. Also in agreement with the earlier findings is even larger variability of maximal *legato* judgments and the small variability of minimal *legato* judgments.
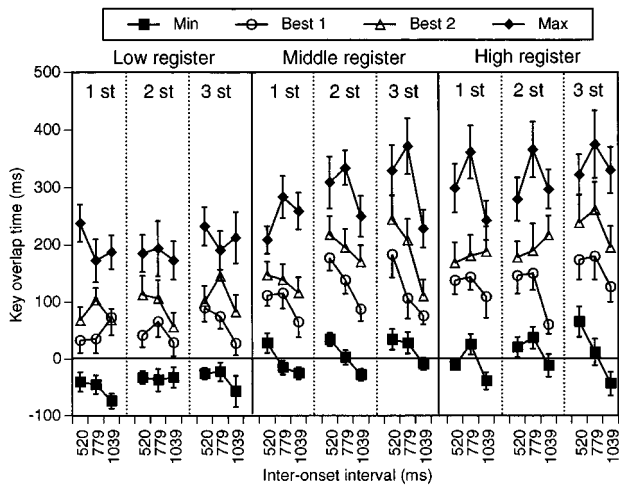
FIG. 6. Perception experiment: Average KOTs and standard error bars for the 27 conditions (3 registers, 3 pitch step sizes, 3 IOIs) in each of four blocks of trials (best 1, min, max, best 2).

Previously, there had also been a clear increase in variability with register, but here there was only a weak tendency in this direction.

There was a pronounced main effect of register in each of the three adjustment conditions, which were subjected to separate ANOVAs. The average KOT for optimal (best) *legato* increased from 73 ms in the second octave to 145 ms in the fourth octave to 169 ms in the sixth octave [$F(2,20) = 17.74$, $p < 0.0001$], and the average KOT for maximal *legato* increased from 199 to 286 to 320 ms [$F(2,20) = 15.48$, $p < 0.0001$]. These increases are in agreement with those found on the digital piano, but the KOTs chosen were a good deal longer here, especially in the low register. Evidently, listeners can tolerate more overlap when they listen to a real piano in a room than when they listen to synthetic piano tones over earphones, despite the much longer postrelease decay times of the realistic piano tones. There was also an effect of register on the KOTs for minimal *legato*, though it only contrasted the second with the fourth and sixth octaves: The average KOT increased from −41 to 5 to 6 ms [$F(2,20) = 17.75$, $p < 0.0001$]. This is also in agreement with the earlier findings, though again the KOTs were longer (i.e., less negative) than before. This is again surprising in view of the slower postrelease decay of the Yamaha tones, which was expected to permit longer nominal gaps in minimal *legato*. On the digital piano, the total range of KOTs that was acceptable as *legato* (max−min) had increased strongly with register; this increase can be seen here, too, but it is less pronounced.

Since the keys above F6 had no dampers, their KOT was not under the listeners' control. Adjustments in the high register (sixth octave) thus derived solely from the two lowest tones in each five-tone sequence. This may account for the relatively small increase in KOTs between the middle and high registers, but it is remarkable that the variability of judgments did not increase substantially in the high register.

The main effect of step size was less pronounced than that of register but nevertheless statistically reliable in each

of the three judgment conditions: Average optimal KOTs increased from 111 ms for chromatic scales to 130 ms for whole-tone scales to 146 ms for diminished-seventh *arpeggi* [$F(2,20) = 5.49$, $p < 0.02$], average maximal KOTs increased from 250 to 266 to 288 ms [$F(2,20) = 4.17$, $p < 0.04$], and average minimal KOTs increased from −22 to −5 and −2 ms [$F(2,20) = 3.85$, $p < 0.04$]. This replicates earlier findings on the digital piano. However, there the effect of step size had decreased in the higher registers. Here this was not apparent. The two-way interaction reached significance only for optimal KOTs [$F(4,40) = 3.05$, $p < 0.03$] but reflected different patterns of the step size effect in different registers, not a decrease in its average size. It is not clear what may have caused those different patterns.

While the preceding effects were fairly straightforward, the effects of IOI (or tempo) were more complex. Average optimal KOTs were 143 ms at the short IOI (fast tempo), 141 ms at the medium IOI or tempo, and 103 ms at the long IOI (slow tempo) [$F(2,20) = 6.84$, $p < 0.006$]. Thus, rather than increasing monotonically, KOT first remained stable and then decreased with tempo. This trend is actually in agreement with the earlier study, where a decrease at the longest IOI was also observed. The present experiment did not include the shortest IOI of the earlier design (260 ms) which had yielded the shortest optimal KOTs. Thus it seems that the relationship between IOI and optimal KOT is nonmonotonic, with a maximum somewhere between 500 and 800 ms of IOI. The results for maximal *legato* suggest that the maximum may be closer to 800 ms. The average maximal KOTs at the three IOIs were 267, 294, and 242 ms, respectively [$F(2,20) = 7.75$, $p < 0.004$]. (The earlier data had not shown a decrease at the longest IOI for maximal *legato*.) The results for minimal *legato* showed a monotonic decrease of the average KOT with increasing IOI, from 8 to −2 to −36 ms [$F(2,20) = 9.47$, $p < 0.002$]. This main effect is in agreement with the earlier findings.

The effect of IOI furthermore interacted with other factors, as had been the case on the digital piano. For optimal *legato*, there were significant interactions of IOI with register [$F(4,40) = 2.93$, $p < 0.04$] and with step size [$F(4,40) = 4.25$, $p < 0.006$]. The first interaction is best characterized as an emergence of a difference between middle and high register as IOI increased. The second interaction can be described either as an increase in the effect of IOI (i.e., in the relative decrease in KOT at the longest IOI) with step size or as a disappearance of the step size effect at the longest IOI. These interactions do not correspond to earlier findings. For maximal *legato*, only the interaction of IOI with register was significant [$F(4,40) = 3.83$, $p < 0.01$]. It, too, suggests an increase in the difference between middle and high registers as IOI increased. For minimal *legato*, the triple interaction between register, step size, and IOI reached significance [$F(8.80) = 2.47$, $p < 0.02$]. However, it does not reflect an orderly pattern.

Finally, it must be mentioned that there were enormous individual differences. Average optimal KOTs for individual listeners ranged from 18 to 189 ms, maximal KOTs from 86 to 378 ms, and minimal KOTs from −78 to 39 ms. The correlation between listeners' average optimal and maximal

KOTs was 0.62 ($p<0.05$), that between their optimal and minimal KOTs was 0.70, and that between their maximal and minimal KOTs was 0.66. Clearly, the listeners differed greatly in ''overlap tolerance,'' and those with high overlap tolerance also had a low ''gap tolerance.'' These subjective criteria seemed to be unrelated to the instruments the participants had learned to play.

On the whole, these results replicate the findings of Repp (1995). The KOT for optimal legato increases with the average pitch (register) of the tones, evidently because the faster decay of high tones reduces acoustic overlap. It increases with step size, presumably due to the increasing consonance of the tones. It increases with IOI initially (Repp, 1995) but then decreases as the IOI approaches 1 s. The initial increase can be understood in terms of decreasing acoustic overlap, but the decrease at longer IOIs must reflect a different factor coming into play. It is not clear at present what that factor might be.

The present data differ from those for the digital piano in that they show greater tolerance for overlap and less tolerance for gaps, a less pronounced widening of the KOT acceptability range as a function of register, and an increase in the preferred KOT in the course of the experiment, as well as some additional differences in detail.

III. PRODUCTION EXPERIMENT

The purpose of the production task was to replicate Repp's (1995) analogous experiment on a different instrument and to re-examine the relationship between perception and production of *legato*. The earlier study had shown that the KOTs do not pattern in exactly the same way in production and in perception. In particular, pianists did not seem to compensate for the effect of register on tone decay, whereas their KOTs were very strongly and monotonically affected by tempo. Only the effect of pitch step size seemed to be comparable in perception and production. Of additional interest were differences between the two hands and patterns of KOTs as a function of position in the tone sequence.

A. Method

1. Pianists

Nine pianists participated whose level of technical skill ranged from advanced amateur to professional. They included two third-year graduate students (artist's diploma candidates), one first-year graduate student, and one second-year undergraduate (performance certificate) student at the Yale School of Music, four Yale University undergraduates who were skilled pianists, and the author, an amateur. Six of the participants were female. The students were paid for their services.

2. Materials and procedure

Thirty-six music sheets were prepared, each showing a sequence of notes in common (4/4) meter. The sequences matched the ones used in the perception study, except that each consisted of exactly three ascending–descending cycles terminated by a longer final note. The nine sequences in the

middle register (fourth octave) were represented twice, with printed instructions to play with either the right or the left hand. The sequences in the low register (second octave) naturally were to be played with the left hand, and those in the high register (sixth octave) with the right hand. Different IOIs were represented by different note values: sixteenth notes, eighth notes, or quarter notes. All sequences were to be played with the fingering 1-2-3-4-5-4-3-2-1 (right hand) or 5-4-3-2-1-2-3-4-5 (left hand). Although this was not necessarily the most preferred fingering for these particular sequences, it was reasonably comfortable at all three pitch step sizes. (This was the reason for centering all sequences on G rather than on some other pitch.)

The pianists were recorded on the Yamaha Disklavier Mark II grand piano from which the acoustic recordings had been made. The piano was connected to a Macintosh computer that recorded the performance in MIDI format. A metronome flashing silently at 60 beats per minute was in view. The pianists were instructed to play at that tempo but in a natural way, not necessarily with mechanical precision. The intended IOIs thus were approximately 250, 500, and 1000 ms, as in Repp's (1995) study. The music pages were shuffled into a different random order for each pianist. Each pianist played the 36 sequences (each of which included three up–down cycles) once.

3. Analysis

Each pianist's MIDI data were imported as text into a spreadsheet program. The different sequences were identified, and their KOTs were calculated by subtracting the onset time of each note from the offset time of the preceding note. Unexpectedly, two of the pianists occasionally held a key down for a very long time (a technique called ''finger pedaling''). To avoid distortion of the data by these exceptionally long KOTs, they were replaced by the average of the preceding and following KOTs. Such replacements amounted to about 1% of the data for each of these two pianists.

For the main statistical analysis, the left- and right-hand data for the middle register were averaged. The KOTs were then entered into a five-way repeated-measures ANOVA with the fixed factors register (3 levels), step size (3 levels), IOI (3 levels), position in the sequence (8 levels), and cycles (3 levels).[10] A four-way ANOVA was conducted on each pianist's individual data, with cycles serving as the random factor. Another ANOVA was conducted on the four pianists for whom hand information was available.[11] In this analysis, the two sets of middle-register data were not averaged, and instead of a single factor of register, two crossed but nonindependent factors were defined: hand/register (2 levels: left-low and left-middle versus right-middle and right-high) and register within hands (2 levels: left-low and right-middle versus left-middle and right-high). Similar analyses were conducted on the four individual pianists, with cycles as the random factor. Their middle-register KOTs were also analyzed separately, contrasting left and right hands with register held constant.
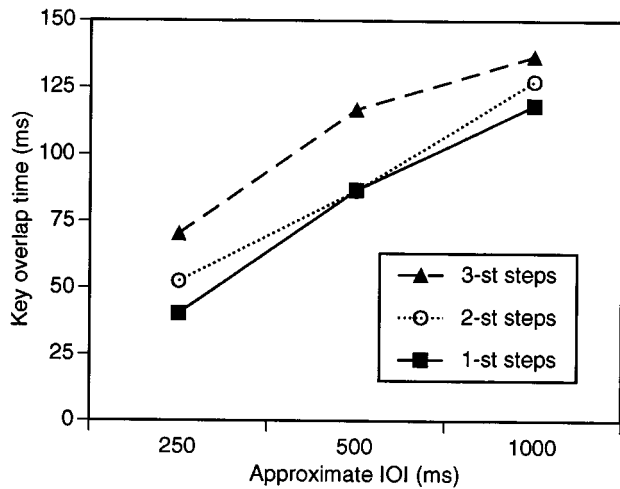
FIG. 7. Production experiment: Average KOTs as a function of (approximate) IOI and step size.
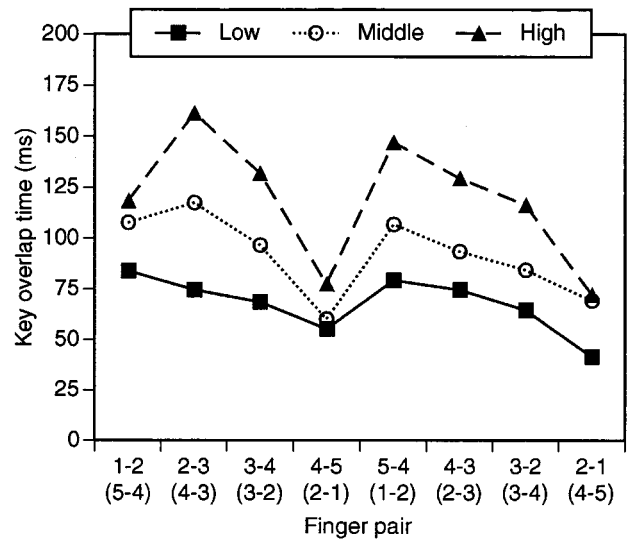

FIG. 8. Production experiment: Average KOTs as a function of position in the sequence (finger pair) and register. Finger pairs in parentheses are for the left hand.

B. Results and discussion

The general range of the average KOTs obtained in the various conditions (about 40 to 140 ms) was similar to that found previously on the digital piano. Despite considerable variability in the raw data, the ANOVAs revealed a relatively simple and highly reliable pattern of results. Figure 7 shows the effects of IOI and step size. The KOTs increased significantly with IOI $[F(2,16) = 12.01, p < 0.001]$. This effect was significant for all nine pianists individually, and it replicates the digital piano data. The KOTs also increased with step size $[F(2,16) = 6.93, p < 0.007]$. This effect was obviously due to the 3-st sequences compared to the other two. All but two individual pianists showed significant step size effects. A similar effect (with 1- and 3-st sequences only) had been found by Repp (1995). The interaction between step size and IOI was not significant.

Two other main effects and an interaction are shown in Fig. 8. KOTs clearly increased with register $[F(2,16) = 11.92, p < 0.001]$. This effect was shown by all pianists but one (the author). Evidently it was not just due to the high register, where the absence of dampers for the higher strings may have invited laziness of the fingers. However, the register difference could have been due to a difference between the left and right hands. In fact, this is the interpretation Repp (1995) proposed for the register effect he found on the digital piano. In the present case, this interpretation seems less plausible. A separate analysis on the four pianists for whom hand information was available showed a significant effect of register within hands $[F(1,3) = 15.01, p < 0.04]$ but not of hand/register $[F(1,3) = 4.23, p < 0.14]$. A separate analysis on the middle-register KOTs likewise did not show a significant hand difference overall $[F(1,3) = 0.71]$. Individually, two pianists had significantly longer KOTs for the right rather than the left hand, one showed a nonsignificant difference in the opposite direction, and one showed no difference. All four, however, showed significant effects of register. This suggests that there was an effect of register apart from any hand differences, although consistent differences in favor of the right hand in the other five pianists could have

contributed to the overall magnitude of the register effect.

The other main effect shown in Fig. 8 is that of position in the sequence or pairs of fingers $[F(7,56) = 7.18, p < 0.0001]$. This pattern, which was shown by all pianists but one, again matches earlier findings on the digital piano: Regardless of the hand used, KOTs decreased in the course of an upward or downward movement on the keyboard and were "reset" to a longer duration after the direction of the sequence had reversed. This effect interacted with register $[F(14,112) = 2.53, p < 0.004]$, mainly because it was more pronounced at higher than at lower pitches. However, it did not interact significantly with step size or, in the separate analysis on four pianists, with hand.

Three additional interactions reached significance in the overall ANOVA. Position interacted with cycles $[F(14,112) = 2.42, p < 0.006]$, due to a tendency towards longer KOTs at the beginning of the first and at the end of the last cycle within a sequence.[12] Cycles were not involved in any other significant effects. The other two significant effects were a triple interaction between position, step size, and register $[F(28,224) = 1.69, p < 0.02]$ and a two-way interaction between register and IOI $[F(4,32) = 4.31, p < 0.007]$. These interactions were primarily due to scaling of effects: For example, effects of register were largest at the longest IOI, and effects of IOI were largest in the high register.

Individual differences in average KOT were very large, with a range from 23 to 185 ms. Short KOTs were shown not only by the author, who as an amateur does not have a refined legato technique, but also by the two most advanced graduate student pianists. The longest KOTs, presumably reflecting the best legato technique, were exhibited by two female undergraduate pianists. A point of considerable interest is that the variability of individual pianists' KOTs increased roughly in proportion to their average KOT; this relationship is shown in Fig. 9. The measure of variability shown is the standard deviation (the root mean square) of the five-way
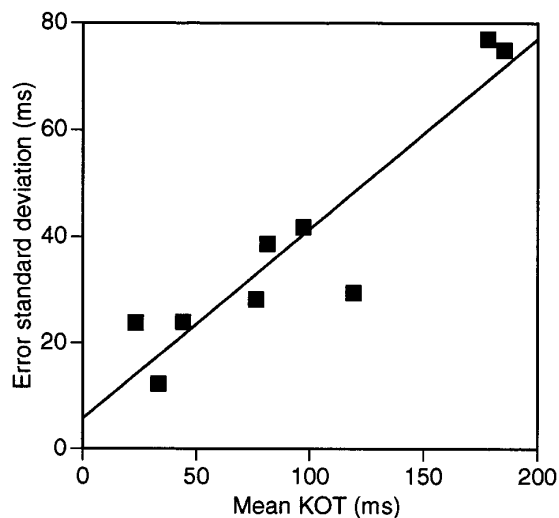
FIG. 9. Variability of KOTs as a function of average KOT duration for individual pianists, with the best-fitting linear regression line.

(cycles×register×IOI×step size×position) interaction in each individual ANOVA. The relationship is reminiscent of analogous findings regarding the precision of IOI timing in finger tapping (e.g., Peters, 1989) and expressive piano performance (Repp, in press), but the KOTs are much shorter than the IOIs for which a generalized Weber's law holds.

Because the variance between cycles was much smaller than that between pianists, many interactions that were nonsignificant in the overall ANOVA reached significance in the analyses on individual pianists' data. However, it would lead too far to discuss these idiosyncratic patterns. One of the two advanced graduate students' data deserve attention, however, because he was the only individual who had also participated in the digital piano study about two years earlier. His average KOT on the digital piano was 34 ms, compared to 23 ms on the Yamaha. His average KOT patterns as a function of position were similar. He showed no effect of step size in either experiment, and his IOI main effect was restricted to an increase in KOT at the slow tempo. In both studies, this pianist showed an interaction between IOI and step size, with a systematic tempo effect only at the 1-st separation. The pattern of his position by step size interaction was also replicated, despite the different black/white key patterns involved in the two studies.[13] Only a paradoxically reversed main effect of register in the earlier study was not replicated here. These observations indicate that individual patterns of KOTs may exhibit some stability over time and across different instruments.

IV. *LEGATO* ARTICULATION IN A SIMPLE MUSICAL PIECE

The present study included the performance of a short piece of music requiring *legato* articulation. The purpose of collecting these additional data was to see whether some of the findings of the preceding experiment would hold up in a simple but meaningful musical context. The music chosen allowed comparison of KOTs between the two hands, although they were confounded with register (about one octave

apart), and comparison between two note values (IOIs). The modulation of KOTs in the course of the music as a function of position in a phrase and of fingering was of interest, as were individual differences among the pianists in average KOT.

A. Method

After playing the scales and *arpeggi* just discussed, the nine pianists sight-read an easy piece requiring *legato* articulation. The composition was *Dialogue*, No. 3 from the *Little Piano Book* of Vincent Persichetti. In this music, shown in Fig. 10, the left and right hands alternate playing a melody consisting of quarter notes and eighth notes while the other hand plays a simple chordal accompaniment. The right-hand melody is almost entirely within the fourth octave; that of the left hand is mostly in the third octave. The melody is divided into seven 2-bar phrases, four for the right hand and three for the left. All but one of these phrases are subdivided by slurs in the original score, but, in order to minimize breaks, the pianists were asked to disregard the original slurs and play each phrase with continuous legato; a single large slur across each phrase was drawn by hand in the score, as shown in Fig. 10.[14] In order to standardize the fingering as much as possible, fingering suggestions were also entered into the score, though it was not possible to verify that they were followed precisely. The suggested tempo was 120 beats per minute, and a silently flashing metronome was in view. The intended IOIs thus were about 250 ms (eighth notes) and 500 ms (quarter notes). Several pianists played at a somewhat slower tempo. Use of the pedal was not permitted. After a brief familiarization with the music, each pianist played the piece three times on the Yamaha Disklavier and was recorded in MIDI format. KOTs were computed for the melodic voice in each phrase, and KOTs at phrase boundaries were ignored.

B. Results and discussion

The KOTs across the positions in the music constitute a ''KOT profile.'' There were very large individual differences among the pianists' profiles, which means that an average KOT profile must be regarded with caution. Nevertheless, short of displaying all individual profiles, it seems most convenient to discuss the data in this way. The average KOT profile is shown in Fig. 11. Rectangles represent the overlaps at the onsets of quarter notes, squares those at the onsets of eighth notes. Dotted vertical lines separate the seven phrases.

The majority of the KOTs (the mode of the distribution of average KOTs, as it were) was in the vicinity of 50 ms. There were a few shorter and a larger number of longer values. The very long KOTs usually represent the average of extremely long KOTs exhibited by some pianists and ''normal'' KOTs shown by others in these positions. Extremely long KOTs reflect ''finger pedaling,'' a key depression that extends during the whole duration of the following note and sometimes beyond. For example, the G4 at the beginnings of bars 11 and 12 was held by some pianists through the following E4. This is a perfectly acceptable technique if the successive notes are consonant, as they are in this case. Finger pedaling was much more common in pianists who had

3
Dialogue

Little Piano Book - 15

FIG. 10. *Dialogue*, No. 3 from the *Little Piano Book* of Vincent Persichetti (© 1954 Elkan-Vogel, Inc.). Reproduced with permission of the publisher (Theodore Presser Company, Bryn Mawr, PA). Long slurs, fingerings, and measure numbers have been added to the score by the author.

relatively long average KOTs, such as the two female undergraduate students mentioned earlier.

Some of the very short KOTs represent grouping within phrases. For example, the very short KOT in the middle of the second phrase occurs at the half-phrase boundary. The shortest KOTs in the fourth and seventh phrases seems to have a different reason, the preparation of a shift in hand position.

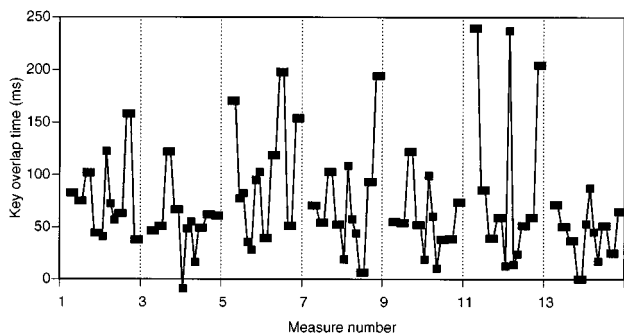Bars 1, 3, 7, and 9 contain similar melodic patterns in

FIG. 11. Average KOTs within the phrases of the Persichetti piece. Squares represent KOTs at the onsets of eighth notes, rectangles KOTs at the onsets of quarter notes.

quarter notes, and they all show similar KOT patterns as well: A longer KOT occurs on the fourth beat (the third data point in Fig. 11), which means that the quarter note on the third beat was prolonged. At the beginnings of bars 2, 4, 8, 10, and 14 there are linearly ascending or descending sequences of four eighth notes. They consistently show a longer KOT for the second note, which means that the first note was prolonged. Both of these tendencies may be consequences of metrical stress: The prolonged notes are in metrically accented positions. To determine the reliability of these observations, separate repeated-measures ANOVAs were performed on these quarter-note and eighth-note data. Both trends were significant [$F(3,24) = 8.74$, $p < 0.001$, and 7.15, $p < 0.002$, respectively], though not every individual pianist exhibited them.

Given the wide range of average KOTs, there does not appear to be a systematic difference between KOTs for eighth and quarter notes, analogous to the effect of IOI in the preceding experiment. Of course, the different IOIs were intermixed here. Nor was there any obvious difference between the right hand (phrases 1, 3, 4, and 6) and the left hand (phrases 2, 5, and 7), at least as far as the shorter KOTs are concerned. This was also true for individual pianists. Very

long KOTs (finger pedaling), however, occurred almost exclusively in the right hand.

Individual differences in average KOT among the pianists were largely preserved: The correlation between their individual average KOTs for tone sequences with IOIs of 500 ms and their median KOTs in the present performances (which seemed preferable over mean KOTs in view of the finger pedaling) was 0.79 ($p < 0.01$).

In general, these data were too variable to reveal simple effects comparable to those observed with the tone sequences. However, the KOT patterns point to structural and ergonomic factors that may come into play in more complex musical materials: phrase boundaries, metrical accents, shifts in hand position, and fingering patterns. These factors deserve more thorough study in future investigations.

V. GENERAL DISCUSSION

A. Piano acoustics

The acoustic analysis part of this study, while mainly serving to characterize the tones used in the subsequent perception and production experiments, does make a modest contribution to the piano acoustics literature which contains relatively little information on the relative sound levels of tones, their decay times (especially after release), and their variability. It should be emphasized that the absolute levels and decay times, as well as their specific patterns of variability as a function of pitch, are specific to the instrument and the recording conditions. However, the data do indicate some general trends and the magnitude of acoustic variability to be expected in a grand piano.

There is a very regular relationship between MIDI velocity and the peak sound level of piano tones, which is well described by a negatively accelerated quadratic function (Fig. 1). Palmer and Brown (1991), in an analysis of tones recorded from a computer-controlled Bösendorfer SE grand piano, found a linear relationship between hammer velocity (in m/s) and peak sound amplitude, which implies a logarithmic relationship between hammer velocity and peak sound level (in dB). The relationship between hammer velocity and MIDI velocity then would also have to be logarithmic in order for the present results to agree with theirs. Indeed, the author found this to be the case on a similar Bösendorfer instrument which had been interfaced with a standard MIDI system, so that both hammer and MIDI velocities could be obtained.[15] If a similar relationship between hammer and MIDI velocities held on the Yamaha Disklavier used here (and there is no particular reason to doubt this, though the hammer velocities were not known), then the present findings are indeed consistent with those of Palmer and Brown.

The present measurements replicate Repp's (1993) finding that the peak sound levels of different piano tones recorded in the same location vary over a range exceeding 10 db, almost half the dynamic range of the instrument (Fig. 2). Importantly, however, there is now also evidence that the specific pattern of these variations varies with the location of the recording microphone.[16] Although careful studies of the perceived loudness of natural piano tones remain to be conducted, it is clear that they do not sound as uneven dynamically as their peak sound levels might suggest, despite the fact that the listener occupies a specific position in space. Although it remains to be shown that a shift in microphone position as small as the distance between a listener's ears can change this pattern of variation significantly, it seems quite possible that binaural hearing and small head movements largely smooth out these unsystematic differences in peak sound level (cf. Benade, 1976, p. 347).

The measurements of the prerelease decay of 1-s piano tones confirmed the well-known fact that high tones decay faster than low tones (Martin, 1947). In addition, they revealed a fairly abrupt discontinuity at the beginning of the fifth octave (C5). Lower tones decayed at a steady rate that increased with pitch, from about 10 to 20 db/s. Higher tones decayed at a much faster rate over the first few hundreds of milliseconds and then at a slower rate beset with irregularities. These irregularities are attributed to beats among slightly mistuned multiple strings (Hundley *et al.*, 1978). The strings of lower tones presumably take longer than 1 s to get out of phase (cf. Martin, 1947), so that the stage of slower and more irregular decay was not reached in the 1-s tones measured here. Martin (1947) has described the two-stage decay of lower tones, but he reported only a single decay rate for higher tones, perhaps because his time resolution was too coarse. Given the observations on string interactions and coupling by Hundley *et al.* (1978) and Weinreich (1990), there is no particular reason why high tones should differ from low tones in that respect; only their strings get out of phase more rapidly, so that the initial fast decay extends only over a few hundred milliseconds. What remains surprising is the abruptness of the change in decay patterns around C5 in the present data.

The acoustic property most pertinent to *legato* articulation, and the one most neglected in previous literature, is the postrelease (damped) decay rate of the piano tones. It was generally between 50 and 120 dB/s, which means that damped tones of high intensity took 0.5 to 1 s to decay completely (i.e., by 60 dB, according to conventional standards). These times probably reflect the reverberation time of the room the instrument was situated in; they may be even longer in a large hall but shorter in an anechoic environment. However, there was also considerable variation from tone to tone, and damped 250-ms tones decayed more quickly than longer tones, for reasons that are not fully understood. It was not determined whether the pattern of variations in decay rate varied with microphone position.

Even though the postrelease decay rates measured here may be room-specific, they represent an important characteristic of the tones heard by the participants in the subsequent experiments. Especially relevant is the fact that the decay rates were considerably slower than those of the digital piano tones used in Repp (1995).

B. Perception of *legato* articulation

Given this difference in postrelease decay rates, it was surprising to find that listeners in the perceptual experiment chose *longer* KOTs in this study than in the previous one. This was true not only for optimal and maximal *legato,* but also for minimal *legato,* which is essentially a gap detection

task. One might have expected gaps to be more difficult to detect when the postrelease decay rate of the tones is slower, but in fact they seemed easier to detect. In some conditions, there was not even a nominal gap (i.e., a negative KOT) associated with minimal *legato* adjustments. Similarly, one might have expected acoustic overlap of successive tones to become noticeable at shorter KOTs when there is a slower and longer postrelease decay. It appears that the perception of connectedness among successive tones is not a simple function of their acoustic contiguity and overlap. Exposure to a realistic piano timbre and room acoustics may have affected listeners' perceptual criteria in the direction of greater tolerance for the sonorities of briefly overlapping tones.

As predicted, listeners' tolerance for acoustic overlap increased as the relative consonance of successive tones increased and as their pitch register increased. The register effect may be due to the increase in decay rate with pitch, but it could also represent a reduction of perceived dissonance among higher tones, due to their fewer and more widely spaced harmonics. To the extent that the second explanation applies, the effect of register may be similar to that of pitch step size. The perceptual roughness accompanying a dissonance may make acoustic overlap more detectable.

The present data, in conjunction with those of Repp (1995), suggest that there is a nonmonotonic relationship between tempo (IOI duration) and the KOT judged to be optimal for legato. Listeners tolerate less overlap at short IOIs (<500 ms) but also show reduced KOTs at long IOIs (>800 ms). This reduction at long IOIs runs counter to the smaller acoustic overlap of long tones, due to their progressive decay. Therefore it must have a different cause. One possibility is that the rhythmic coherence of successive tones begins to be weakened at IOIs of 1000 ms, whose rate is almost twice as slow as that of the preferred pulse or tactus (Parncutt, 1994). Perceptual integration of the ''heads'' and ''tails'' of decaying tones into unitary temporal forms and consequent perceptual segregation of overlapping tones seems to underlie listeners' remarkable tolerance for (and indeed, failure to perceive) simultaneities extending over hundreds of milliseconds. At long IOIs, this temporal integration may begin to weaken, so the acoustic overlap of the ''tails'' of tones with the ''heads'' of following tones may become more noticeable.

C. Production of *legato* articulation

A comparison of the present production data with those of Repp (1995) would seem to be relevant to the issue of pianists' adjustment to different instruments and acoustic environments. Precise comparisons are difficult because of the different individuals involved and the large individual differences in average KOT within each group. Nevertheless, the finding that the average KOTs were similar on the real piano and on the digital piano, even though the postrelease decay times differed by a factor of 2, suggests that *legato* technique is not very flexible in response to extramusical factors. While small adjustments (detectable only in a within-pianist comparison) may well occur, pianists' general motor habits seem to be fairly well entrenched. Indeed, it is very difficult for someone who typically plays legato with short KOTs (such

as the author) to produce long KOTs, and while pianists who have acquired this *legatissimo* technique through long practice may be able to play with short KOTs, they are unlikely to abandon the fruit of their hard labor unless the music specifically requires it. Basically, it seems that a characteristic individual *legato* style is maintained regardless of the specific acoustic conditions.

The effects of tempo and step size on KOTs replicate earlier findings on the digital piano. The tempo effect is likely to have a kinematic cause (increased general tension and/or earlier preparation for finger lifting at faster tempi), especially since the increase in KOTs from medium to slow tempo contradicts the decrease observed in the perceptual task. The step size effect, too, could be kinematic in origin (spread of fingers), though it is tempting to conclude that its cause is aesthetic, due to relative consonance. While the direction of the effect agrees with that found in perception, there was little difference in production between the 1-st and 2-st conditions, which clearly differed in perception. The reason for this discrepancy is not clear.

Only the effect of register was free of kinematic influences, though it was partially confounded with possible hand differences which, unfortunately, could be assessed in only some of the pianists. Nevertheless, there was an indication of an independent effect of register on KOTs which parallels a similar effect in perception and may reflect a compensatory adjustment to the more rapid decay of high tones. In addition, some pianists seem to play somewhat more *legato* with the right than with the left hand, and ''finger pedaling'' occurred almost exclusively in the right hand.

One of the most consistent findings on the digital piano, the dependence of KOT on position in the tonal sequence, was fully replicated here. It suggests that KOT is a contextually modulated parameter that may carry structural and expressive information within the *legato* range. This result is the one most deserving of follow-up investigation with more varied musical materials. A small step in that direction was taken here by examining the KOTs in performances of Persichetti's *Dialogue*, in which unexpectedly large variability of KOTs was encountered, due in part to finger pedaling by some pianists. Future studies may have to employ somewhat more controlled materials in order to determine more precisely the structural and ergonomic factors that influence KOT. One likely factor is the distance between keys in conjunction with fingering, and future research on *legato* articulation may dovetail nicely with ongoing work on pianists' fingering strategies (e.g., Parncutt *et al.*, in press).

[1]This is sometimes referred to as finger legato, since the same effect can be achieved by holding and releasing the first tone with the damper pedal. The present study does not deal with legato achieved through pedaling.

[2]A baby grand piano of the same make was used in the perception experiment. The change in pianos was beyond the author's control. Acoustic recordings were made from the baby grand piano as well but only partially analyzed, as they were found to contain some distortion.

[3]Nominal durations are given as specified in the MIDI instructions. Due to a peculiarity of the MAX software, however, the actual durations were 3.9% longer than specified.

[4]Decay times are usually given in terms of the time it takes for the sound level to drop by 60 dB. However, the dynamic range of the present recordings was only about 40 dB, so that conventional decay times could not be measured directly.

[5]The difference in absolute sound level between the instruments is meaningless.

[6]Repp (1995) found a decrease in the peak intensity of digital piano tones in the highest (sixth) octave, but this may have been an artifact of using a fixed 30-ms integration window.

[7]On the present Yamaha Disklavier, C2–G2 were double-stringed, while all pitches from G#2 on were triple-stringed.

[8]These are the actual durations. The intended intervals had been multiples of 250 ms, but their actual durations turned out to be 3.9% longer, due to the MAX software.

[9]As pointed out earlier, the later replacement of the instrument was not under the author's control. Since the lid of the piano had been removed by other users and could not easily be put back, the experiment was run without the lid on. The acoustic conditions thus were not identical with those described in the previous section: Both the instrument and the listeners' distance from it were different. Arguably, however, the acoustic characteristics of the tones heard by the listeners were more similar to those of the Yamaha grand piano than to those of the digital piano in Repp's (1995) study.

[10]Repp (1995) called the position factor ''pairs of fingers.'' The pairs of fingers are 1-2, 2-3, 3-4, 4-5, 5-4, 4-3, 3-2, 2-1 for the right hand and 5-4, 4-3, 3-2, 2-1, 1-2, 2-3, 3-4, 4-5 for the left hand; that is, 1-2 in the right hand is aligned with 5-4 in the left hand, and so on. Therefore ''position in the sequence'' seems a better name for this factor.

[11]In order to be able to determine later whether a middle-register sequence had been played with the right or the left hand, it was necessary to record the random sequence in which the music sheets were presented to each pianist. Unfortunately, this step was omitted for the first five pianists. This oversight was embarrassing but not fatal, as hand differences were only of secondary interest.

[12]There may also have been longer IOIs in these positions, reflecting traces of expressive timing. (No analysis of IOI durations was conducted.)

[13]While the 1-st sequences were comparable (WBWBW; W=white, B=black), the 3-st sequence was WBBWW on the digital piano but BW-WBB on the Yamaha.

[14]One pianist, contrary to instructions, followed the slurs in the original score and thus showed large negative KOTs (i.e., breaks) at several points within phrases. These KOT values were excluded from the data, as were a few isolated instances of large negative KOTs in other pianists' data. The pianist was also excluded from the analysis of variance.

[15]The precise function was $MV = 52.3 + 45.7 \ln(HV)$, where MV = MIDI velocity and HV = hammer velocity. The relationship between hammer velocity and peak sound level was also logarithmic. (Unpublished data collected in the course of a study by Repp, 1993.) It is not obvious why peak sound level should be a quadratic function of MIDI velocity, but quadratic functions fit the data in Fig. 1 distinctly better than do logarithmic functions.

[16]This was also found in similar measurements conducted on tones recorded from the baby grand piano used in the perception experiment.

Benade, A. H. (**1976**). *Fundamentals of Musical Acoustics* (Oxford U.P., New York; reprinted in 1990 by Dover, New York).

Hundley, T. C., Benioff, H., and Martin, D. W. (**1978**). ''Factors contributing to the multiple rate of piano tone decay,'' J. Acoust. Soc. Am. **64,** 1303–1309.

Kuwano, S., Namba, S., Yamasaki, T., and Nishiyama, K. (**1994**). ''Impression of smoothness of a sound stream in relation to legato in musical performance,'' Percept. Psychophys. **56,** 173–182.

Martin, D. W. (**1947**). ''Decay rates of piano tones,'' J. Acoust. Soc. Am. **19,** 535–541.

Palmer, C., and Brown, J. C. (**1991**). ''Investigations in the amplitude of sounded piano tones,'' J. Acoust. Soc. Am. **90,** 60–66.

Parncutt, R. (**1994**). ''A perceptual model of pulse salience and metrical accent in musical rhythms,'' Music Percept. **11**, 409–464.

Parncutt, R., Sloboda, J. A., Clarke, E. F., Raekallio, M., and Desain, P. (**in press**). ''An ergonomic model of keyboard fingering for melodic fragments,'' Music Percept.

Peters, M. (**1989**). ''The relationship between variability of intertap intervals and interval duration,'' Psychol. Res. **51**, 38–42.

Repp, B. H. (**1993**). ''Some empirical observations on sound level properties of recorded piano tones,'' J. Acoust. Soc. Am. **93,** 1136–1144.

Repp, B. H. (**1995**). ''Acoustics, perception, and production of *legato* articulation on a digital piano,'' J. Acoust. Soc. Am. **97,** 3862–3874.

Repp, B. H. (**in press**). ''Variability of timing in expressive piano performance increases with interval duration,'' Psychonom. Bull. Rev.

Weinreich, G. (**1990**). ''The coupled motion of piano strings,'' in *Five Lectures on the Acoustics of the Piano*, edited by A. Askenfelt (Royal Swedish Academy of Music, Stockholm), pp. 73–81.

Perception of synthetic /ba/–/wa/ speech continuum by budgerigars (*Melopsittacus undulatus*)

Micheal L. Dent, Elizabeth F. Brittan-Powell, Robert J. Dooling, and Alisa Pierce
Department of Psychology, University of Maryland, College Park, Maryland 20742

Other than humans, extensive vocal learning has only been widely demonstrated in birds. Moreover, there are only a handful of avian species that are known to be good mimics of human speech. One such species is the budgerigar (*Melopsittacus undulatus*), which is a popular mimic of human speech and learns new vocalizations throughout adult life. Using operant conditioning procedures with a repeating background task, we tested budgerigars on the discrimination of tokens from two synthetic /ba/–/wa/ speech continua that differed in syllable, but not transition, duration. Budgerigars showed a significant improvement in discrimination performance on both continua near the phonetic boundary for humans. Budgerigars also showed a shift in the location of the phonetic boundary with a change in syllable length, similar to what has been described for humans and other primates. These results on a nonmammalian species provide support for the operation of a general, nonphonetic, auditory process as one mechanism which can lead to the well-known stimulus-length effect in humans. © *1997 Acoustical Society of America.* [S0001-4966(97)03409-7]

PACS numbers: 43.80.Lb, 43.71.Es, 43.66.Gf [FD]

INTRODUCTION

Budgerigars (*Melopsittacus undulatus*) are small Australian parrots that have recently been shown to produce a variety of human speech phonemes either singly, during warble-song, or in response to presented objects (Banta and Pepperberg, 1995). These birds have also proven to be excellent subjects for psychoacoustic studies. Recent studies have shown that budgerigars, like humans, perceive the vowel tokens /i/, /a/, /e/, and /u/ in phonetically appropriate categories in spite of variation in talker, pitch contour, and gender (Dooling and Brown, 1990; Dooling, 1992). In other studies on the perception of consonants by budgerigars (Dooling *et al.*, 1989), perceptual boundaries were near the human boundaries for voice-onset-time (VOT) pairs of /ba/–/pa/ (bilabial), /da/–/ta/ (alveolar), and /ga/–/ka/ (velar) continua. Perception of speech sound categories and discrimination among phonetically relevant speech tokens are not unique to budgerigars. Other studies have shown that zebra finches, starlings, quail, blackbirds, and pigeons also discriminate and categorize speech sounds similar to humans (Hienz *et al.*, 1981; Kluender, 1991; Dooling, 1992; Dooling *et al.*, 1995).

There are several reasons why it is interesting to study how birds discriminate among speech sounds. First, some species of birds such as starlings, mynahs, and some parrots modify their vocalizations throughout their lives and are so flexible in what they will accept as acoustic models for learning that they even mimic human speech (e.g., Thorpe, 1959; Greenewalt, 1968; Klatt and Stefanski, 1974; Dooling, 1986; Pepperberg, 1990; Farabaugh *et al.*, 1994; Patterson and Pepperberg, 1994; Warren *et al.*, 1996). These well-known instances of human speech production show that birds can extract the important acoustic features of speech, despite different tutors. Evidence from a variety of studies shows they can imitate many phonemes, including all vowel sounds (Patterson and Pepperberg, 1994) and almost every conso-

nant sound in the English language (Turney *et al.*, 1994). Second, birds that can produce speech potentially provide the only known opportunity for testing the relation between production and perception of human speech in a nonhuman organism. Third, it is well known that budgerigars and other birds have the remarkable ability to regenerate their hair cells following acoustic overexposure or treatment with ototoxic drugs (Cotanche, 1987; Corwin and Cotanche, 1988; Ryals and Rubel, 1988). In the case of budgerigars, both the perception of species-specific vocal signals and the precision in production of learned contact calls are affected when hearing is lost from ototoxic drugs. However, both production and perception of these vocal signals recover as the auditory periphery becomes repopulated with new hair cells (Dooling *et al.*, in press). Because of this regenerative capability, birds provide a unique model for studying speech perception (and production in the case of budgerigars) after destruction and subsequent repair of the peripheral auditory system. Fourth, experiments with various speech categories and speech continua have shown that mammals such as chinchillas and monkeys with auditory capabilities similar to humans tend to show phonetic boundaries and categories similar to the human phonetic boundaries (Burdick and Miller, 1975; Kuhl and Miller, 1975; Kuhl and Padden, 1982; Kuhl, 1987). Birds have both peripheral and central auditory systems that are profoundly different than the mammalian auditory system (for a review, see Manley, 1990; Carr, 1992; Manley and Gleich, 1992), so they contribute a different perspective on the role of mammalian auditory processing in speech perception.

Thus there are numerous reasons for studying speech perception in birds in addition to the general comparative strategy of using animals to test whether a particular perceptual performance exhibited by humans listening to speech is a uniquely human phenomenon. A recent study by Sinnott and her colleagues (Sinnott *et al.*, submitted) is intriguing in

this respect and provides some of the motivation for the present study. They show that a nonhuman primate exhibits a boundary along a human speech continuum that is usually explained, in humans, by normalization for articulation rates. These results are also interesting because nonhuman mammals lack a supralaryngeal cavity and are incapable of producing human speech (for a review see Ploog, 1992). Birds, on the other hand, do not use their larynx for sound production. Instead, the syrinx acts as the sound source and is positioned such that birds have a suprasyringeal cavity. Several species have been shown to mimic human speech sounds and phrases extremely well.

Many experiments with humans have investigated the consonant–vowel (CV) /ba/–/wa/ continuum and have shown that information occurring later in the speech stream affects the perception of an earlier occurring cue (Miller and Liberman, 1979; Godfrey and Millay, 1981; Pisoni *et al.*, 1983; Diehl and Walsh, 1989). As syllable duration increases, the /b/–/w/ perceptual boundary for humans moves toward transitions of longer duration, resulting in a phonetic boundary shift. Though it is difficult to conceive of a purely psychoacoustic explanation for this stimulus-length effect (see, for example, Pisoni *et al.*, 1983; Diehl and Walsh, 1989), comparative work tends to support psychoacoustic explanations for perceptual categories, in general.

One psychoacoustic explanation that has been suggested as playing a role in the shift of boundary location in the stimulus-length effect is backward masking (Jamieson, 1987). When the steady-state portion of the CV syllable increases, the transition may be more effectively masked, leading to a boundary shift. By showing that reducing the vowel intensity level also shifts boundary location, Jamieson provided evidence that some form of backward masking may provide the basis for such an effect.

In earlier experiments, quail performed similar to humans in their discrimination of several phonetic categories, and the authors argue that phonetic categories are natural auditory groupings, even though speech has a different functional significance for humans and nonhumans (Diehl and Kluender, 1989). These authors (Diehl and Kluender, 1989) suggest an auditory enhancement hypothesis for speech perception, where human listeners are particularly sensitive to the auditory cues that define phonetic categories and may typically discriminate changes at phonetic boundaries more efficiently than animals.

A more prevalent explanation for these results is the phonetic one that suggests listeners are exhibiting perceptual normalization for speaking. Miller and Liberman (1979) suggest that the listener interprets a set of acoustic cues in running speech in relation to the speaker's rate of articulation rather than by reference to some absolute value, a theory supported by many other studies on speech perception (e.g., Dorman *et al.*, 1977; Gay, 1978; Repp *et al.*, 1978; Repp, 1982). When Miller and Liberman added a stop consonant to the end of their syllables (/bad/–/wad/), the phonetic boundary shift occurred in the opposite direction, suggesting that listeners normalize for articulation rate.

Because we know that budgerigars can produce /b/ and /w/, the present experiment sought to examine how budgeri-
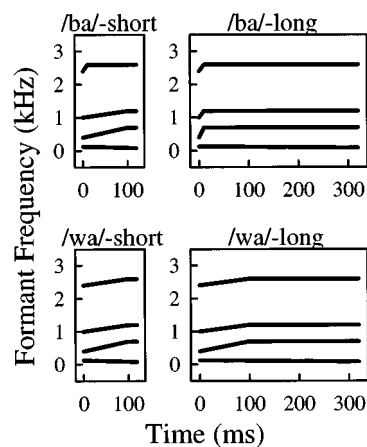

FIG. 1. Schematic of the spectrogram of the end points of the /ba/ and /wa/ continua for the short (120 ms) and long (320 ms) stimulus sets.

gars discriminated among tokens on a synthetic /ba/–/wa/ speech continuum. For these tests, we used a standard discrimination (low-uncertainty) task that has been used in a variety of other experiments on the discrimination of both simple and complex sounds in these birds (Dooling *et al.*, 1987). While budgerigars have demonstrated perceptual boundaries for other consonant contrasts using discrimination procedures, the /ba/–/wa/ contrast affords the opportunity to test the phenomenon of the boundary shift with increases in syllable duration.

I. METHODS

A. Subjects

Three adult budgerigars (two females and one male) were used as subjects. All of the birds were housed in a vivarium at the University of Maryland and were kept on a day/night cycle corresponding to the season. The birds were either purchased from a local pet store or bred in the vivarium. They were kept at approximately 90% of their free-feeding weight during the course of the experiment. In addition, four adult humans who were native speakers of English were used as subjects. All four were female students, ranging in age from 24 to 25 yr, working in the laboratory at the time of the experiment but who had no previous testing experience with these sounds. None of the subjects reported a history of speech or hearing disorders or spoke other languages fluently.

B. Stimuli

The stimuli used in this experiment were two full-formant /ba/–/wa/ speech continua differing in length (Fig. 1). These speech sounds were generated by the Canadian speech research environment (CSRE) program (using a 10-kHz sampling rate) according to the parameters of Sinnott *et al.* (submitted). The duration of the initial formant transition changed in 10-ms steps to yield a set of ten stimuli ranging perceptually from /ba/ to /wa/. Figure 1 shows a schematic representation of the end-point stimuli for both the short and long continua. For both continua, $F1$ began at 400 Hz and moved to 700 Hz over a variable time period ranging

from a 10-ms /ba/ to a 100-ms /wa/. $F2$ moved from 1000 to 1200 Hz, and $F3$ moved from 2400 to 2600 Hz over the same time periods. $F4$ and $F5$ had no formant transitions. $F0$ fell linearly over the duration of the syllable from 125 to 80 Hz. All stimuli were gated externally with a 10-ms rise time. The duration of the steady-state vowel was manipulated so that the overall duration of the short syllables was 120 ms and the long syllables were 320 ms.

We chose to use these stimuli instead of the other stimulus continua on which humans have been tested (Miller and Liberman, 1979; Godfrey and Millay, 1981; Pisoni *et al.*, 1983; Diehl and Walsh, 1989) to facilitate comparison with the only other animal data available on this speech continuum. All stimuli were presented at a peak sound-pressure level of 68 dB.

C. Apparatus and procedure

The apparatus and procedure for testing the birds has been described previously (Okanoya and Dooling, 1987, 1991). The response panel and procedures are shown schematically in Fig. 2. The birds were tested in a wire cage ($23 \times 25 \times 16$ cm) that was placed in a small, foam-lined, sound isolation booth (IAC model IAC-2; $57 \times 60 \times 78$ cm). A response panel consisting of two sensitive microswitches with light-emitting diodes (LEDs) was mounted on the wall of the test cage just above the food hopper. The microswitch was tripped by the bird pecking the LED. The left microswitch and LED served as the observation key, and the right microswitch and LED served as the report key. The speech stimuli were delivered from a JBL loudspeaker (model 2105H) mounted above the test cage. The experiment was controlled by an IBM 486 microcomputer operating Tucker–Davis Technology electronic and DSP modules. The behavior of the animals during test sessions was monitored continuously by a video camera system.

The birds were trained by a standard operant autoshaping program to peck the observation key during a continuously repeating background (interstimulus-interval of 380 ms) and to peck the report key when a new sound (target) was presented alternately with the background sound. A peck to the report key within 2 s following onset of an alternating sound pattern was considered a hit. A hit was reinforced with a 2-s access to food on a schedule of 80%–100%. Stimuli from the long CV continuum were tested first, and stimuli from the short CV stimuli were tested second.

During the testing phase, a peck on the observation key began a random interval of 1–6 s. Following this interval, the next peck on the observation key initiated a trial (defined as an alternation of the target stimulus with the background stimulus). The dependent variable in these experiments was the response latency on each trial involving a target stimulus (i.e., the alternating sound pattern with the target stimulus alternating with the background stimulus). Previous work on the perception of both simple and complex sounds (including speech sounds) has shown that response latency is a valid measure of stimulus similarity for these birds (Dooling *et al.*, 1995). If the bird failed to respond (miss), a response latency of 2000 ms was recorded, and a new trial sequence began with the observation interval for the next trial. Approxi-
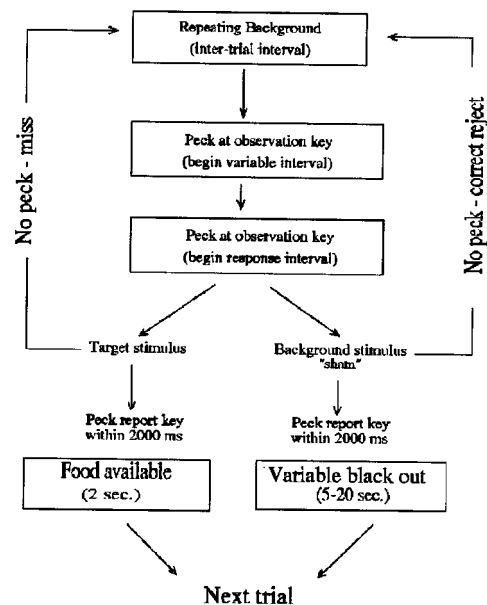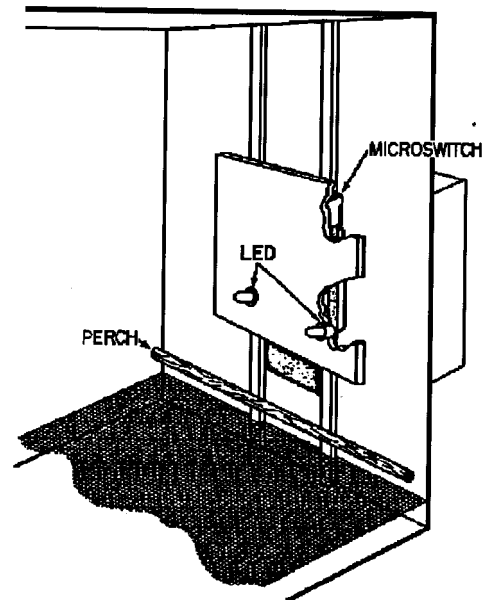

FIG. 2. Schematic of testing apparatus showing the arrangement of LEDs, food hopper, and speaker (top) and a diagram of the alternating sound task (bottom). After the variable interval begins, the next peck on the observation key starts the response interval. During the response interval, one of two trial types occurs: (1) if a target stimulus is presented to the bird, pecking the report key yields a food reward or (2) if a sham stimulus is presented, a peck to the report key causes a blackout interval. Failure to peck the report key during each of the trials starts a new intertrial interval.

mately 10% of the trials were sham trials in which the target stimulus was the same as the background stimulus. A response on the report key during a sham trial or during the waiting interval was punished with a variable blackout period during which the lights in the chamber were extinguished while the repeating background continued. The blackout time ranged from 5–20 s depending on the bird's recent history for false responses.

The strategy in this experiment was to test each stimulus in the continuum against every other stimulus in the con-

tinuum. The order of stimulus testing was determined prior to the experiment. A matrix of stimuli (background×target) was constructed, and one row was selected at random from this matrix for testing. The first stimulus in the row then served as the background (the continuously repeating stimulus), and the remaining stimuli in the row served as targets (selected randomly on a trial-by-trial basis). Testing continued until each target had been tested ten times and all rows were tested (i.e., until each sound served both as a background and as a target stimulus ten times). At the conclusion of testing (usually ten sessions), a 10×10 matrix of response latencies was available for analysis where each cell in the matrix contained a mean response latency based on the ten trials of each stimulus pair.

To ensure that these stimuli were satisfactory replicas of natural speech, we also tested human subjects on the same stimuli using two procedures: the discrimination task the birds were tested on and a labeling task. In the discrimination task, humans listened to the stimuli through headphones (Realistic model Nova 67) while pressing response keys on a hand-held panel. The human subjects viewed a video screen showing the inside of the animal operant chamber to provide feedback as to whether their response was correct (a hit, hopper activation) or incorrect (false alarm, lights extinguished). In addition, humans subjects were also tested on an identification task in which they were required to label each stimulus played to them as either /ba/ or /wa/. In this test, the subjects pressed a button on a computer keyboard to randomly play one of the ten tokens, and then wrote down their response. In all, the ten tokens from each continuum were played ten times in random order, for a total of 100 trials for each continuum. Results were analyzed as the percent of responses labeled /ba/ for each subject, and the boundary along each continuum was defined as the 50% point.

D. Analysis

At the conclusion of testing on the repeating background discrimination task, matrices of percent correct values and response latencies were obtained for each subject (budgerigars and humans) for each stimulus contrast along the continua (Dooling *et al.*, 1995). The 10×10 matrices described above were folded about the diagonal, and the corresponding cells in the upper and lower halves were averaged to obtain triangular matrices that contained an average value from 20 trials for each stimulus contrast for each subject (10 vs 20 combined with 20 vs 10). From these matrices, we obtained both percent correct data and response latency data for each subject for all possible pairwise contrasts. A two-way analysis of variance (ANOVA) was used to test whether response latencies between the short and long continua differed, and whether the specific stimulus pairs along the continua differed from each other. The stimulus comparisons were done for both one-step contrasts (e.g., 10/20) and two-step contrasts (e.g., 10/30).

II. RESULTS

The one-step stimulus comparisons were almost impossible, with each bird averaging less than 20% correct. There-
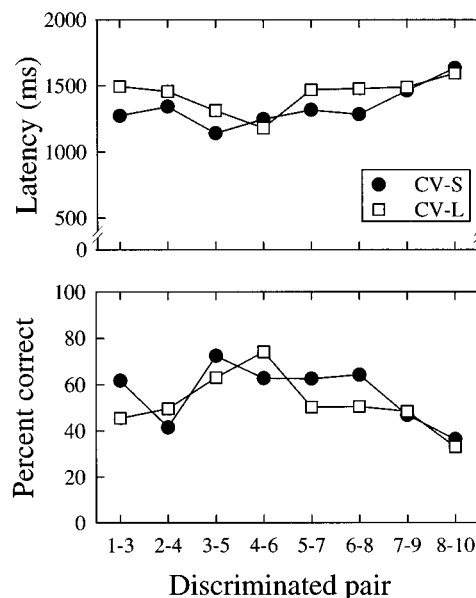

FIG. 3. Discrimination performance for the budgerigars for the two step stimuli comparison. Both response latency (top panel) and percent correct (bottom panel) are plotted as a function of length and show a peak in the discrimination function that shifts to the right with the increase in syllable duration.

fore we analyzed only two-step stimulus comparisons, which the birds discriminated among more easily for both the short and long stimulus conditions (Fig. 3). The two-step discrimination task response latency was inversely correlated with percent correct for all three birds both on the short ($r = -0.98, -0.90, -0.98$) and long ($r = -0.99, -0.99, -0.89$) stimulus continua. For the short stimulus (CV-S) continuum, the three budgerigars showed the best discrimination for the 30/50 stimulus pair (percent correct=72, response latency=1140 ms while for the long stimulus (CV-L), the 40/60 pair was discriminated best (percent correct=74, response latency=1180 ms).

Across both continua, budgerigars were significantly better at discriminating the CV-S continuum than the CV-L continuum [$F(1,7) = 5.29$, $p < 0.05$]. There were also significant differences along both the continua [$F(7,7) = 4.68$, $p < 0.05$], with discrimination performance better near the center of each continuum and with the discrimination peak in a different location for short compared to long duration syllables (30/50 for CV-S and 40/60 for CV-L).

Humans were tested on both a repeating background discrimination task as well as an identification task. Like the birds, they discriminated among tokens in the CV-S continuum more easily than those of the CV-L continuum [$F(1,7) = 8.85$, $p < 0.05$]. There were also significant differences along the continua [$F(7,7) = 6.86$, $p < 0.05$]. When tested on the identification task, humans showed clear phonetic boundaries and boundary shifts similar to those typically reported by speech researchers using other /ba/–/wa/ continua. Humans showed slightly higher boundary locations than the budgerigars but a similar boundary shift with increase in stimulus duration. For the CV-S continuum, the boundary was 49 ms, but for the CV-L continuum, the
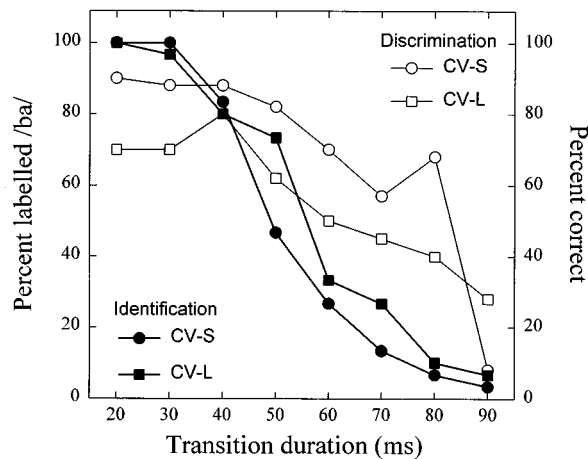
FIG. 4. Discrimination and identification performance for the humans for the two-step stimuli comparisons. Percent correct for the discrimination task (open symbols) show decreases in performance with the increase in transition duration (from /ba/ to /wa/). The boundaries for the identification task (closed symbols) shift from 49 ms for the short to 56 ms for the long stimuli.

boundary was 56 ms. The discrimination and identification results for humans are shown in Fig. 4.

The shift in human boundaries from the identification task were similar to shifts in the peaks in the discrimination functions in the CV-S and CV-L continua for budgerigars. However, the peaks in discrimination for the human subjects were not nearly as clear as for the birds. Instead, human performance showed a change in discrimination that followed Weber's law (percent correct decreased and response latency increased as transition durations increased), an effect noted by others when humans are tested on some speech sound discriminations with a low uncertainty task.

III. DISCUSSION

It is well known that budgerigars can mimic a variety of human speech sounds. Despite the fact that budgerigar peripheral (Manley, 1990; Manley and Gleich, 1992) and central (Brauth et al., 1987; Brauth, 1988; Striedter, 1994) auditory systems are remarkably different from the mammalian auditory system, budgerigars have nevertheless been shown to discriminate among a variety of speech sounds including /ba/–/pa/, /ga/–/ka/, /da/–/ta/, /ra/–/la/ and a variety of natural and synthetic vowels (Dooling et al., 1989, 1995; Dooling and Brown, 1990). The present results add to this database by showing that budgerigars discriminate among /ba/–/wa/, phonemes they are capable of producing (Turney et al., 1994), in a way that is consistent with the stimulus-length effect shown in humans.

Studies of speech perception in birds, particularly those that are speech mimics, may have particular relevance for understanding the evolution of human acoustic communication. Several recent studies of parrot phonation have shown that parrots may use their tongue to produce vowels in some ways which are similar to humans (Patterson and Pepperberg, 1994; Warren et al., 1996) and that tracheal resonances may affect certain aspects of vocal production (Brittan-Powell et al., 1997). Moreover, some psittacines also demonstrate a remarkable ability to mimic a variety of nonvocal

human behaviors, including arm, leg, and head movements (Moore, 1992). We know from both anecdotal as well as scientific evidence that budgerigars show socially dependent vocal flexibility throughout life (Farabaugh et al., 1994; Banta and Pepperberg, 1995; Brittan-Powell et al., in press). Taken together, these findings raise the possibility that exposure to talking humans may sensitize these birds to some of the critical acoustic features of human speech. Budgerigars, for instance, show a peak in discrimination of a sinewave /ra/–/la/ continuum whereas zebra finches tested under identical conditions do not (Dooling et al., 1995).

A wealth of comparative studies on the perception of speech by animals have shown that uniquely human structures and perceptual mechanisms are not necessary for obtaining humanlike discrimination and classification of English speech sounds (e.g., Kuhl and Miller, 1975; Hienz et al., 1981; Dooling et al., 1989, 1995; Kluender, 1991; Sinnott et al., submitted). Birds provide an important addition to this database both because their auditory system is significantly different from that of humans and other mammals and because they have the ability to regenerate auditory hair cells following acoustic overexposure or treatment with ototoxic drugs (Cotanche, 1987; Corwin and Cotanche, 1988; Ryals and Rubel, 1988). Elsewhere, we have shown that following hair cell damage, both perception and production of species-typical contact calls are only disrupted for a short time before returning to normal (Dooling et al., in press). These vocal signals recover as the auditory periphery becomes repopulated with new hair cells (Dooling et al., in press), so these birds provide a unique model for studying speech perception after repair of the peripheral auditory system.

It is also worthwhile considering whether there is a common mechanism producing the stimulus-length effect in human, nonhuman primates, and budgerigars. At least in budgerigars and nonhuman primates, one nonphonetic explanation for the stimulus-length effect is backward masking, where later information in the syllable masks earlier important cues. This may be true as well of humans. This common acoustic explanation based on backward masking would be strengthened if there was evidence for similar backward masking thresholds in budgerigars and humans for nonspeech stimuli. Fortunately, nonsimultaneous auditory masking data are available for budgerigars, and we know that they show backward (but not simultaneous or forward) masking thresholds that are nearly identical to those described in humans (Dooling and Searcy, 1980). To some extent then, the similar shift in the /ba/–/wa/ phonetic labeling boundaries for humans and the shift in the location of the discrimination peak in budgerigars may be due to similar sensitivities to backward masking. Such results would be consistent with the auditory account of the stimulus-length effect in the perception of /ba/–/wa/ as proposed by Diehl and Walsh (1989), who found such effects not only with speech continua but also with nonspeech stimuli.

It is important to recognize that there are well-known effects of testing paradigm on speech discrimination and

speech perception findings including the stimulus-length effect in the /ba/–/wa/ continuum. Phonetic boundaries tend to follow Weber's law with weak or nonexistent phonetic boundaries when a low uncertainty discrimination task was used compared to when a high uncertainty identification task was used (Kewley-Port *et al.*, 1988). MacMillan *et al.* (1988) found similar results with vowels, but not consonants. Our human subjects showed clear labeling phonetic boundaries, but no corresponding peaks in their discrimination functions. When tested on an identification task our human subjects showed not only the expected phonetic boundaries for the /ba/–/wa/ continua, but also the expected shift with increasing syllable length. A review of the literature shows that as syllable duration increases from about 100 ms to about 300 ms, the boundary changes from around 30 ms to around 45 ms (Miller and Liberman, 1979; Godfrey and Millay, 1981; Pisoni *et al.*, 1983; Diehl and Walsh, 1989; Sinnott *et al.*, submitted). The budgerigars in this study had peaks in discrimination functions at 40 ms for the 120-ms stimuli and 50 ms for the 320-ms stimuli. This shift falls in the range of the data from humans and monkeys, suggesting common mechanisms.

One popular view of the biology of speech suggests that the auditory system has driven the selection of speech sound categories by providing natural perceptual categories, or broad acoustic targets, for speech sounds that the articulatory system has subsequently evolved to match. In the course of language acquisition, these broad perceptual categories or proclivities may be altered by environmental input during development to become more precisely defined (for a review, see Kuhl, 1989). This leads to, among other things, the many well-known cross cultural differences in the production and perception of speech sound categories (see, for example, Miyawaki *et al.*, 1975). At the other end of the continuum are arguments for phonetic, uniquely human specializations for the perception of speech, such as the findings that even young infants correctly perceive a particular consonant such as /d/ even when the acoustic information for the consonant /d/ varies tremendously depending on speaker, speaking rate, and vowel context, among other things (for a review, see Kuhl, 1989).

The present results contribute to what we know of the biology of speech by providing more evidence that ''general'' properties of the vertebrate auditory system underlie the perception of speech sound categories—in this case a bird, the budgerigar, shows the vowel length effect. Changing the rate of speaking obviously results in a number of changes in the acoustics of the speech stream, including eliminating and reducing pauses between words and phrases and shortening words. Shortening words brings with it changes in the temporal and spectral cues affecting both vowels and consonants. It seems reasonable to suppose that speakers may vary some features of speech (e.g., vowel duration) over other features when changing speaking rate because the auditory system demands it. Changes in some features may be necessary to maintain optimal processing of other, more critical features. An animal example of such a strategy comes from the horseshoe bat (*Rhinolophus ferrumequinum*). These bats compensate for flight-induced Doppler shifts in the frequency of their echoes by lowering the frequency of subsequent calls (Schnitzler, 1973). In this way, the frequency of the echo returning to the bat remains in a narrowly tuned frequency range which is disproportionally represented throughout the auditory system from the cochlea to the cortex (Neuweiler, 1980; Pollak, 1980). This specialization allows the horseshoe bat to optimize the processing of the returning echoes.

There are clearly aspects of speech perception that cannot be easily explained by the simple auditory mechanisms as described above, which argue for other hypotheses. For the present speech contrast, the phonetic interpretation argues that human listeners use rate-based articulation cues to adjust their perception of phonemes. A compelling case for this argument is seen in the work of Miller and Liberman (1979) where the addition of a stop consonant on the end of the syllable produced a boundary shift in the opposite direction. It would be interesting to conduct a corresponding experiment in budgerigars and perhaps other birds as one way of testing the generality of auditory accounts of the stimulus-length effect. The point is not to dispute the fact that linguistic experience is important for general phonetic categorizations in humans. Instead, the present data showing that budgerigars exhibit the stimulus-length effect seen in humans, and the comparative approach to speech perception in general, aims to push the auditory account of the stimulus-length effect to its limit. To the extent this can be done, then uniquely common human capabilities are not *required* for perceiving the sounds of speech.

ACKNOWLEDGMENTS

Banta, P. A., and Pepperberg, I. M. (**1995**). ''Learned English vocalizations as a model for studying budgerigar (*Melopsittacus undulatus*) warble song,'' in *Nervous Systems and Behavior Proceedings of the 4th International Congress of Neuroethology*, edited by M. Burrows, T. Matheson, P. L. Newland, and H. Schuppe (Thieme, New York), p. 335.

Brauth, S. E., and McHale, C. M. (**1988**). ''Auditory pathways in the budgerigar. II. Intratelencephanic pathways,'' Brain Behav. Evol. **32**, 193–207.

Brauth, S. E., McHale, C. M., Brasher, C. A., and Dooling, R. J. (**1987**). ''Auditory pathways in the budgerigar. I. Thalamo-telencephalic pathways,'' Brain Behav. Evol. **30**, 174–199.

Brittan-Powell, E. F., Dooling, R. J., and Farabaugh, S. M. (**in press**). ''Vocal development in budgerigars (*Melopsittacus undulatus*): Contact calls,'' J. Comp. Psych.

Brittan-Powell, E. F., Dooling, R. J., Larsen, O. N., and Heaton, J. T. (**1997**). ''Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*),'' J. Acoust. Soc. Am. **101,** 578–589.

Burdick, C. K., and Miller, J. D. (**1975**). ''Speech perception by the chinchilla: discrimination of sustained /a/ and /i/,'' J. Acoust. Soc. Am. **58,** 415–427.

Carr, C. E. (**1992**). ''Evolution of the central auditory system in reptiles and birds,'' in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 511–543.

Corwin, J. T., and Cotanche, D. A. (**1988**). ''Regeneration of sensory hair cells after acoustic trauma,'' Science **240,** 1772–1774.

Cotanche, D. A. (**1987**). ''Regeneration of hair cell stereociliary bundles in the chick cochlea following severe acoustic trauma,'' Hearing Res. **30,** 181–196.

Diehl, R. L., and Kluender, K. R. (**1989**). ''On the objects of speech perception,'' Ecol. Psych. **1**, 121–144.

Diehl, R. L., and Walsh, M. A. (**1989**). ''An auditory basis for the stimulus-length effect in the perception of stops and glides,'' J. Acoust. Soc. Am. **85**, 2154–2164.

Dooling, R. J. (**1986**). ''Perception of vocal signals by budgerigars (*Melopsittacus undulatus*),'' Exp. Biol. **45**, 195–218.

Dooling, R. J. (**1992**). ''Perception of speech sounds by birds,'' in *Advances in Biosciences: Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, London), pp. 407–413.

Dooling, R. J., Best, C. T., and Brown, S. D. (**1995**). ''Discrimination of synthetic full-formant and sinewave /ra–la/ continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*),'' J. Acoust. Soc. Am. **97**, 1839–1846.

Dooling, R. J., and Brown, S. D. (**1990**). ''Speech perception by budgerigars (*Melopsittacus undulatus*): Spoken vowels,'' Percept. Psychophys. **47**, 568–574.

Dooling, R. J., Manabe, K., and Ryals, B. M. (**1996**). ''Effect of masking and hearing loss on vocal production and vocal learning in budgerigars,'' Assoc. Res. Otolaryngology Abstr. 581.

Dooling, R. J., Okanoya, K., and Brown, S. D. (**1989**). ''Speech perception by budgerigars (*Melopsittacus undulatus*): The voiced-voiceless distinction,'' Percept. Psychophys. **46**, 65–71.

Dooling, R. J., Park, T. J., Brown, S. D., Okanoya, K., and Soli, S. D. (**1987**). ''Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*). II. Vocal signals,'' J. Comp. Psych. **101**, 367–381.

Dooling, R. J., Ryals, B. M., and Manabe, K. (**in press**) ''Recovery of hearing and vocal behavior after hair cell regeneration,'' Proc. Natl. Acad. Sci.

Dooling, R. J., and Searcy, M. H. (**1980**). ''Forward and backward auditory masking in the parakeet (*Melopsittacus undulatus*),'' Hearing Res. **3**, 279–284.

Dorman, M. F., Studdert-Kennedy, M., and Raphael, L. J. (**1977**). ''Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues,'' Percept. Psychophys. **22**, 109–122.

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.

Gay, T. (**1978**). ''Effect of speaking rate on vowel formant transitions,'' J. Acoust. Soc. Am. **63**, 223–230.

Godfrey, J. J., and Millay, K. K. (**1981**). ''Discrimination of the tempo of frequency change cue,'' J. Acoust. Soc. Am. **69**, 1446–1448.

Greenewalt, C. H. (**1968**). *Bird song: Acoustics and Physiology* (Smithsonian, Washington, DC).

Hienz, R. D., Sachs, M. B., and Sinnott, J. M. (**1981**). ''Discrimination of steady-state vowels by blackbirds and pigeons,'' J. Acoust. Soc. Am. **70**, 699–706.

Jamieson, D. G. (**1987**). ''Studies of possible psychoacoustic factors underlying speech perception,'' in *The Psychophysics of Speech Perception NATO Advanced Research Workshop on Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Kluwer Academic, Hingham, MA), pp. 220–229.

Kewley-Port, D., Watson, C. S., and Foyle, D. C. (**1988**). ''Auditory temporal acuity in relation to category boundaries; speech and nonspeech stimuli,'' J. Acoust. Soc. Am. **83**, 1133–1145.

Klatt, D. H., and Stefanski, R. A. (**1974**). ''How does a mynah bird imitate human speech?,'' J. Acoust. Soc. Am. **55**, 822–832.

Kluender, K. R. (**1991**). ''Effects of first formant onset properties on voicing judgments result from processes not specific to humans,'' J. Acoust. Soc. Am. **90**, 83–96.

Kuhl, P. K. (**1987**). ''The special-mechanisms debate in speech research: Categorization tests on animals and infants,'' in *Categorical Perception: The Groundwork of Cognition*, edited by S. Harnad (Cambridge U.P., Cambridge, England), pp. 355–386.

Kuhl, P. K. (**1989**). ''On babies, birds, modules, and mechanisms: A comparative approach to the acquisition of vocal communication,'' in *The Comparative Psychology of Audition*, edited by R. J. Dooling and S. H. Hulse (Lawrence Erlbaum, Hillsdale, NJ), pp. 379–419.

Kuhl, P. K., and Miller, J. D. (**1975**). ''Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants,'' Science **190**, 69–72.

Kuhl, P. K., and Padden, D. M. (**1982**). ''Enhanced discrimination at the phonetic boundaries for the voicing feature in macaques,'' Percept. Psychophys. **32**, 542–550.

Macmillan, N. A., Goldberg, R. F., and Braida, L. D. (**1988**). ''Resolution for speech sounds: Basic sensitivity and context memory on vowel and consonant continua,'' J. Acoust. Soc. Am. **84**, 1262–1280.

Manley, G. A. (**1990**). *Peripheral Hearing Mechanisms in Reptiles and Birds* (Springer-Verlag, Berlin).

Manley, G. A., and Gleich, O. (**1992**). ''Evolution and specialization of function in the avian auditory periphery,'' in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 561–580.

Miller, J. L., and Liberman, A. M. (**1979**). ''Some effects of later-occurring information on the perception of stop consonant and semivowel,'' Percept. Psychophys. **25**, 457–465.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. (**1975**). ''An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English,'' Percept. Psychophys. **18**, 331–340.

Moore, B. R. (**1992**). ''Avian movement imitation and a new form of mimicry: Tracing the evolution of a complex form of learning,'' Behav. **122**, 231–263.

Neuweiler, G. (**1980**). ''Auditory processing of echoes: Peripheral processing,'' in *Animal Sonar Systems*, edited by R. J. Busnel and J. F. Fish (Plenum, New York), pp. 519–548.

Okanoya, K., and Dooling, R. J. (**1987**). ''Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds,'' J. Comp. Psych. **101**, 7–15.

Okanoya, K., and Dooling, R. J. (**1991**). ''Perception of distance calls by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Poephila guttata*): Assessing species-specific advantages,'' J. Comp. Psych. **105**, 60–72.

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**, 635–648.

Pepperberg, I. M. (**1990**). ''Some cognitive capacities of an African Gray parrot,'' in *Advances in the Study of Behavior*, edited by P. J. B. Slater, J. S. Rosenblatt, and C. Beer (Academic, New York), pp. 357–409.

Pisoni, D. B., Carrell, T. D., and Gans, S. J. (**1983**). ''Perception of the duration of rapid spectrum changes in speech and nonspeech signals,'' Percept. Psychophys. **34**, 314–322.

Ploog, D. W. (**1992**). ''Evolution of vocal communication,'' in *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, edited by H. Papousek, U. Jurgens, and M. Papousek (Cambridge U.P., New York), pp. 6–30.

Pollak, G. D. (**1980**). ''Organizational and encoding features of single neurons in the inferior colliculus of bats,'' in *Animal Sonar Systems*, edited by R. J. Busnel and J. F. Fish (Plenum, New York), pp. 549–587.

Repp, B. H. (**1982**). ''Perceptual integration and differentiation of spectral cues for intervocalic stop consonants,'' Percept. Psychophys. **24**, 471–485.

Repp, B. H., Liberman, A. M., Eccardt, T., and Pesetsky, D. (**1978**). ''Perceptual integration of acoustic cues for stop, fricative, and affricative manner,'' J. Exp. Psychol. Hum. Percept. Perform. **4**, 621–637.

Ryals, B. M., and Rubel, E. W. (**1988**). ''Hair cell regeneration after acoustic trauma in adult Corturnix quail,'' Science **240**, 1774–1776.

Schnitzler, H. U. (**1973**). ''Control of doppler shift compensation in the Greater Horseshoe bat, *Rhinolophus ferrumequinum*,'' J. Comp. Physiol. **82**, 79–92.

Sinnott, J. M., Brown, C. H., and Borneman, M. A. (**submitted**). ''Effects of syllable duration on stop-glide identification in syllable-initial and syllable-final position by humans and monkeys.''

Striedter, G. S. (**1994**). ''The vocal control pathways in budgerigars differ from those in songbirds,'' J. Comp. Neurol. **343**, 35–56.

Thorpe, W. H. (**1959**). ''Talking birds and the mode of action of the vocal apparatus of birds,'' Proc. Zool. Lond. **132**, 441–455.

Turney, S. M., Banta, P. A., and Pepperberg, I. M. (**1994**). ''Comparative acoustical analyses of learned English vocalization of two parrot species,'' Anim. Beh. Soc. Abstr. 140.

Warren, D. K., Patterson, D. K., and Pepperberg, I. M. (**1996**). ''Mechanisms of American English vowel production in a grey parrot (*Psittacus erithacus*),'' The Auk **113**, 41–58.

LETTERS TO THE EDITOR

Comments on "Effect of multilayer baffles and domes on hydrophone response" [J. Acoust. Soc. Am. 99, 1883–1893 (1996)]

Evgeni L. Shenderov

Research and Development Institute "Morfizpribor," 46 Chkalovski pr., St. Petersburg, 197376, Russia

In the subject paper Ebenezer and Abraham used the transfer matrix method to investigate the response of a receiver located between a layered dome and a layered baffle. They considered this set as one multilayer system and obtained complicated matrix equations. The purpose here is to show that the sound field and the response of the receiver may be expressed as simple expressions using reflection and transmission coefficients having clear physical meaning. © *1997 Acoustical Society of America.* [S0001-4966(97)03209-8]

In the subject paper[1] Ebenezer and Abraham investigated the response of the hydrophones located in the gap between a planar baffle and a planar dome, Fig. 1(a). Each of these is a multilayered set. The authors considered this system as a single multilayered system containing n_1+n_2+3 layers, where n_1 and n_2 are the numbers of layers in the dome and in the baffle, respectively, 3 comes from the lower and the upper half-spaces and the gap. They started this problem from scratch, i.e., derived the velocities and the stresses in all layers, matched boundary conditions, and used the transfer matrix method. They obtained complicated matrix expressions. Very interesting and useful results are presented in the plots.

The goal of this note is to show that the problem may be solved by a much more simple method. We can unify all layers of the dome and all layers of the baffler in two separate systems as shown in Fig. 1(b). Each complex layer is specified by its reflection and transmission coefficients. We can consider the sound field in the gap as an infinite sum of partial waves reflected from the dome and the baffle. This method is described in Refs. 2 and 3 for three-layer systems. The partial waves propagating in the positive z direction are expressed as

$$p_0 B_d A_b (A_d A_b)^{n-1} \exp[ik_z(h+2d(n-1))-ik_x x],$$

where p_0 is the incident wave magnitude, $k_x = k \sin\theta$, $k_z = k \cos\theta$, B_d is the transmission coefficient of the dome, A_d and A_b are the reflection coefficients for the dome and the baffle, respectively, n is the number of the wave ($n=1,...,\infty$), k is the wave number in water, d is the distance between the dome and the baffle, and the time dependence is $e^{-i\omega t}$. The complex coefficients B_d, A_d, and A_b depend on the angle of

incidence of the sound wave. The partial waves propagating in the negative z direction are

$$p_0 B_d (A_d A_b)^{n-1} \exp[ik_z(-h+2d(n-1))-ik_x x]. \quad (1)$$

After summation of the geometrical progressions over n we obtain the sound pressure in the gap

$$p(x,h) = p_0 B_d \exp(-ik_x x)$$
$$\times [\exp(-ik_z h) + A_b \exp(ik_z h)]/C, \quad (2)$$

where $C = 1 - A_b A_d \exp(i2k_z d)$.

This formula is an exact expression valid for an arbitrarily layered dome and baffle. The reflection and transmission coefficients may be calculated by any method, for example, using transfer matrices. In some cases much simpler expressions for these coefficients may be used.

The denominator C describes resonance phenomena in the gap. The expression in the brackets describes a standing wave near the baffle.

The response for an acoustically transparent lengthy receiver located in the gap parallel to the x axis is a product of the response of the receiver in the free field and of the expression Eq. (2) without the factor $\exp(-ik_x x)$.

In the same manner we can solve the other problem considered in Ref. 1 to obtain the sound pressure in the gap generated by the baffle rear side vibrating with the specified velocity

$$W_2 = W_0 \exp(-i\gamma_0 x). \quad (3)$$

Consider an auxiliary plane-wave incident upon the system from the lower half-space (see Fig. 2)

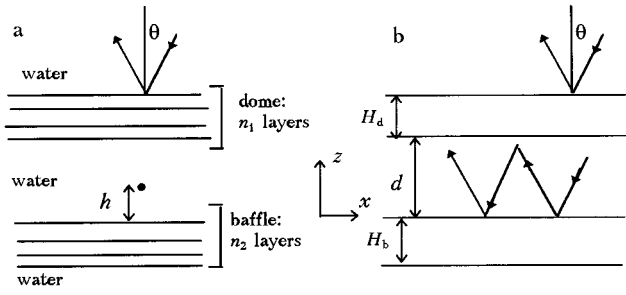$$p_i = p_0 \exp(-i\gamma_0 x + i\alpha_0 z), \quad (4)$$

FIG. 1. The multilayered planar baffle and the dome. (a) Multilayered system, (b) equivalent three-layered system.

where $\alpha_0 = \sqrt{k^2 - \gamma_0^2}$. The origin in Fig. 2 is chosen on the lower side. The partial waves propagating in the gap in the positive and negative directions may be written as

$$p_0 B_b (A_d A_b)^{n-1} \exp[i\alpha_0(h + 2d(n-1) + H_b) - i\gamma_0 x], \quad (5a)$$

$$p_0 A_d B_b (A_d A_b)^{n-1} \exp[i\alpha_0(-h + 2dn + H_b) - i\gamma_0 x], \quad (5b)$$

respectively, where H_b is the total thickness of the baffle. After summation of the geometrical progressions over n we obtain the sound pressure in the gap

$$p_{\text{gap}}(x,h) = p_0 B_b \exp(i\alpha_0 H_b - i\gamma_0 x)$$
$$\times [\exp(i\alpha_0 h) + A_d \exp(i\alpha_0(2d - h))]/C. \quad (6)$$

The total sound field in the lower half-space is a sum of the incident wave, the wave reflected from the baffle and the partial waves Eq. (5b) transmitted through the baffle once more, i.e., of the waves
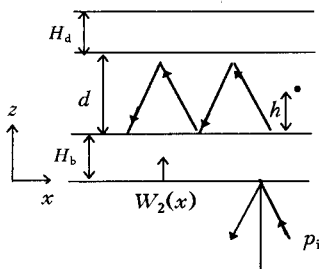

FIG. 2. Radiation of sound by the baffle rear side with the velocity distribution $W_2(x)$.

$$p_0 A_d B_b^2 (A_d A_b)^{n-1} \exp[i\alpha_0(2dn + 2H_b - z) - i\gamma_0 x]. \quad (7)$$

After summation over n we obtain the total sound field in the lower half-space

$$p_{\text{total}} = p_0 \exp(-i\gamma_0 x)[\exp(i\alpha_0 z) + A_{1b} \exp(-i\alpha_0 z)$$
$$+ A_d B_b^2 \exp(i\alpha_0(2H_b + 2d - z))/C], \quad (8)$$

where A_{1b} is the reflection coefficient of the baffle when the sound wave is incident upon the rear side of the baffle. Note that for nonsymmetrical baffles the values A_b and A_{1b} may be different.

The normal component of the velocity of the lower side of the baffle is

$$W_2 = \frac{1}{i\omega\rho} \left. \frac{\partial p_{\text{total}}}{\partial z} \right|_{z=0}, \quad (9)$$

where ρ is the density of water. From Eqs. (6)–(9) we obtain the relation between the sound pressure in the gap and the velocity of the baffle rear side

$$\frac{p_{\text{gap}}}{W_2}$$
$$= \frac{\omega\rho B_b}{\alpha_0} \frac{\exp(i\alpha_0 H_b)[\exp(i\alpha_0 h) + A_d \exp(i\alpha(2d - h))]}{(1 - A_{1b})C - A_d B_b^2 \exp(2i\alpha_0(d + H_b))}. \quad (10)$$

For the system without the dome we put $A_d = 0$, $C = 1$ and obtain a very simple expression

$$\frac{p_{\text{gap}}}{W_2} = \frac{\omega\rho}{\alpha_0} \frac{B_b}{1 - A_{1b}} \exp(i\alpha_0(H_b + h)). \quad (11)$$

Despite the fact that the reflection and transmission coefficients of the baffle are present in Eqs. (10) and (11), the ratio p_{gap}/W_2 does not depend on parameters of the liquid in the lower half-space.

I am grateful to Dr. D. D. Ebenezer for interesting discussions of some acoustical problems.

[1] D. D. Ebenezer and P. Abraham, ''Effect of multilayer baffles and domes on hydrophone response,'' J. Acoust. Soc. Am. **99**, 1883–1893 (1996).
[2] E. L. Shenderov, *Wave Problems in Hydroacoustics* (Sudostroenie, Leningrad, 1972) [English translation: Joint Publications Research service, Report No. JPRS-58146 (433 pp.), 1973. National Technical Information service, Springfield, VA 22151].
[3] I. S. Hickmann, D. E. Risty, and E. S. Stewart, ''Properties of sandwich-type structures as acoustic window,'' J. Acoust. Soc. Am. **29**, 858–868 (1957).

On the eigenfrequency of a gas bubble in a liquid

Johan L. Leander

Department of Military Technology, National Defence College, P.O. Box 27805, S-11593 Stockholm, Sweden

This investigation concerns free linear gas bubble oscillations in liquids. Of prime interest is the eigenfrequency, and in particular its real part, here named as the transient frequency. The conceptual difference between the more frequently consulted resonance frequency and the transient frequency is first addressed by means of the classical mechanical oscillator. Next, bubble pulsations in liquids are discussed and an existing model is used for the gas–liquid interaction from which an approximate expression for the eigenfrequency is derived. A rationale for the approximate evaluation of the functions modeling the thermal processes is suggested which is independent of the frequency content of any possible pressure excitation. Moreover, compressibility effects are not approximated in the derivation presented here. The quantitative difference between the adiabatic resonance frequency and the derived estimate of the transient frequency is found to be of significance for small bubbles. Finally, the similarity between a standard mechanical oscillator and a bubble in a liquid for the case of liquid-compressibility effects only is discussed. © *1997 Acoustical Society of America.* [S0001-4966(97)02409-0]

INTRODUCTION AND REVIEW

Acoustics in gas–liquid suspensions has fascinated man for a long time.[1–7] It has been of general interest to understand the propagation of small-amplitude as well as finite-amplitude waves in a bubbly mixture. The presence of bubble clouds in the oceans and their acoustic effects have resulted in numerous theoretical and experimental investigations concerning, for example, their scattering and absorption properties.[8–9] Furthermore, in recent years, the problem concerning the bubble contribution to underwater ambient noise has been frequently addressed.[10–12] Also, steady-state conditions seem to have rendered most attention. However, and with no exception, an understanding of the above problems relies in one way or the other on the one bubble problem. That is, the interaction between a single gas bubble and the surrounding liquid.[13,14]

In this work we look at free linear oscillations of a single gas bubble in a liquid. Of prime concern is the eigenfrequency of the bubble. That is, the complex quantity involving information concerning the decay of the free bubble pulsations as well as their transient frequency. Perhaps we should first note that strictly speaking, the frequency concept is only applicable for periodic undamped motion. However, there is no need here for an academic attitude, and the quantities to be discussed are easy to visualize and interpret. In order to study the above we first rehash some fundamentals of the classical mechanical oscillator and the corresponding equation of motion is here given in the form[15,16]

$$\Delta \ddot{r} + 2\beta \Delta \dot{r} + \omega_0^2 \Delta r = f_e, \tag{1}$$

where β, ω_0, and f_e are the damping constant, the undamped resonance frequency, and the excitation, respectively. Δr is the bubble radius displacement and the dot represents derivation with respect to time. For a monochromatic forcing with complex amplitude F_e we find from Eq. (1),

$$\frac{\Delta R}{F_e} = \frac{1}{\omega_0^2 - \omega^2 + i2\omega\beta}, \tag{2}$$

where ΔR is the complex amplitude of the radius displacement, ω is the angular driving frequency, and i is the imaginary unit. Maximum response occurs for a minimum of the magnitude of the denominator in the right-hand side of Eq. (2). This particular frequency, the damped resonance frequency, is given by $\omega_{0d} = \sqrt{\omega_0^2 - 2\beta^2}$. Let us now turn to free pulsations and look at the solution of the homogeneous form of Eq. (1) where the forcing is set to zero. The result is

$$\Delta r = A e^{-\beta_{\text{free}} t} \sin(\omega_{\text{free}} t + \psi), \tag{3}$$

where the real constants A and ψ are governed by the initial conditions.[15,16] The complex eigenfrequency is $\omega_{\text{eig}} = \omega_{\text{free}} + i\beta_{\text{free}}$ where $\omega_{\text{free}} = \sqrt{\omega_0^2 - \beta^2}$ is the transient frequency and $\beta_{\text{free}} = \beta$. From the two separate solutions discussed above, we have two conceptually different frequency quantities where the first, the damped resonance frequency ω_{0d}, is applicable for harmonic oscillations and the second, the transient frequency ω_{free}, is applicable for free oscillations. For the academic case of no damping, they coincide, both qualitatively and quantitatively.

In the following sections we focus on transient oscillations and derive an approximate expression for the eigenfrequency by means of an existing model for bubble pulsations in liquids. In doing that, we introduce a new method for approximation of the quantities related to the thermal processes. Also, in the current procedure, compressibility effects are not approximated. We then compare our result with the adiabatic resonance frequency which is frequently used for bubble size estimation.[10,11] It is found that the adiabatic resonance frequency overestimates the transient frequency, which is particularly true for small bubbles. Finally we address the case of compressibility effects, only for which the

mechanical oscillator analogy can be given up front, although the approximate bubble resonance frequency and damping function are frequency dependent for the case of harmonic excitation.

I. THE BUBBLE MODEL

The theoretical foundation of bubble dynamics is based on the fundamental conservation and balance theorems of continuum mechanics (details are given elsewhere).[5,13] Here, we briefly rehash the information needed for our purpose. Let us consider a two phase system consisting of a spherical gas bubble in a compressible liquid that extends to infinity. The bubble is insonified by an acoustic pressure wave given by $P_e \exp(i\omega t)$ where the wavelength is large compared to the equilibrium bubble radius R_0. The bubble radius is decomposed into $R_0 + \Delta R \exp(i\omega t)$, and linear motion is considered, so we assume that the radial displacement is small compared to the bubble radius, $|\Delta R| \ll R_0$. It can then be shown that ΔR is related to P_e via an equation which is analogous to Eq. (2)

$$\frac{\Delta R}{P_e} = \frac{-Q}{\rho_0 R_0}, \tag{4}$$

where Q is of the form

$$Q = \frac{1}{\omega_b^2 - \omega^2 + 2i\omega\beta_{tot}}. \tag{5}$$

The quantity ρ_0 is the liquid density. For the approximate bubble resonance frequency $\omega_b^2 = \omega_t^2 + \omega_r^2$, and the damping function $\beta_{tot} = \beta_\nu + \beta_t + \beta_r$, we have

$$\omega_t^2 = \frac{P_{b0}}{\rho_0 R_0^2}\left[\text{Re }\varphi - \frac{2\sigma}{R_0 P_{b0}}\right], \quad \omega_r^2 = \frac{(\omega k_0 R_0)^2}{1 + (k_0 R_0)^2}, \tag{6}$$

and

$$\beta_\nu = \frac{2\mu}{\rho_0 R_0^2}, \quad \beta_t = \frac{P_{b0}\text{ Im }\varphi}{2\rho_0 R_0^2 \omega}, \quad \beta_r = \frac{\omega^2 R_0}{2c_0[1 + (k_0 R_0)^2]}. \tag{7}$$

The shear viscosity of the liquid is denoted by μ, and $k_0 = \omega_0/c_0$ is the wave number where c_0 is the corresponding speed of sound. The quantities β_ν, β_t, and β_r are the damping functions for shear viscosity, thermal effects, and radiation, respectively. The equilibrium pressure in the bubble is given by

$$P_{b0} = P_0 + \frac{2\sigma}{R_0}, \tag{8}$$

where σ is the surface tension and P_0 is the ambient pressure in the liquid. The function φ is

$$\varphi = \frac{3\gamma}{1 - 3(\gamma-1)i\chi[\sqrt{i/\chi}\text{ coth}(\sqrt{i/\chi}) - 1]}, \tag{9}$$

where $\chi = D/\omega R_0^2$ and D is the thermal diffusivity of the gas. The quantity γ is the ratio for the specific heats of the gas. We consider in the next section free oscillations on the basis of the model presented above.

II. FREE BUBBLE PULSATIONS

We will now discuss free bubble oscillations and in particular the corresponding transient frequency. This will be carried out by means of Fourier theory, and the complex amplitudes from the previous section, ΔR and P_e, will from here on be interpreted as Fourier transforms. However, for convenience, let us first shortly relate some relevant results for a standard mechanical oscillator for the case of a sineburst driver. Thus, an oscillator initially at rest (zero displacement and velocity) is suddenly excited by a sinusoidal excitation of finite duration. By solving this classical problem analytically one finds that for times within the duration of the driver, the response is given by two contributions.[15,16] The first is a decaying oscillation with a frequency corresponding to the transient frequency discussed above. The second part is a pure sinusoidal motion with the same frequency as the driver. Now, if the time scales are given so that the first part becomes negligible within the extent of the duration of the excitation, the system has reached steady-state conditions which thus are governed by the second part of the solution. Further, once the excitation has stopped, the response is given by one expression which is analogous to the first contribution discussed above. In summary, the part that we are looking for here can thus in principle be found in two ways: at the beginning of the motion that takes place when the excitation has just started, or at the beginning of the motion when the excitation has just stopped.

We now set up the relevant Fourier integral corresponding to Eq. (4) and start to look for the free pulsations that will eventually yield an expression for the eigenfrequency. So, given a transient excitation as $p_e(t)$ at the bubble liquid interface, the corresponding radius displacement, $\Delta r(t)$, can formally be found from the relation

$$\Delta r(t) = \frac{-1}{2\pi\rho_0 R_0} \int_{-\infty}^{\infty} P_e(\omega)Q(\omega)\exp(i\omega t)d\omega, \tag{10}$$

where $P_e(\omega)$ is the Fourier transform of $p_e(t)$. The result we are looking for, the eigenfrequency, is governed by the system parameters, damping, etc., and is thus to be independent of the excitation. This means that we are essentially free to choose the excitation. Moreover, since we are about to invert the transform analytically, it is reasonable to use a simple $P_e(\omega)$. We here take the academic but fruitful dirac pulse where $P_e(\omega) = 1$. Now, due to the involved expressions for $\omega_t(\omega)$ and $\beta_t(\omega)$, Eqs. (6), (7), and (9), it is found that the quantity $Q(\omega)$ is also involved, and analytical inversion of the integral Eq. (10) is difficult, if not impossible. However, if one could find approximate but relevant expressions for $\omega_t(\omega)$ and $\beta_t(\omega)$ it would perhaps be possible to invert Eq. (10) analytically. It has been suggested by Prosperetti *et al.* that if only one frequency can be used for the evaluation of $\omega_t(\omega)$ and $\beta_t(\omega)$, the rationale is as follows; the resonance frequency should be evaluated for the driving frequency, $\omega_t = \omega_t(\omega_{drive})$, and the thermal damping function should be evaluated for the corresponding ω_t, $\beta_t = \beta_t[\omega_t(\omega_{drive})]$.[13] We here make the observation that the above suggested procedure is dependent upon the driving frequency. Thus, this is of no use for us and we have to look for a new rationale.

Fortunately, however, this is easily done. In keeping the compressibility effects untouched it is convenient to rewrite $Q(\omega)$ as

$$Q(\omega) = \frac{1 + ik_0 R_0}{\omega_t^2 - \omega^2[1 + 2R_0(\beta_\nu + \beta_t)/c_0] + i\omega[2(\beta_\nu + \beta_t) + \omega_t^2 R_0/c_0]}. \tag{11}$$

Although the quantity β_r is not seen explicitly in Eq. (11), compressibility effects are indeed there (c_0 is there and finite!). Now, if $\omega_t(\omega)$ and $\beta_t(\omega)$ can be properly approximated by two constants, the quantity $Q(\omega)$ will be a rationale and the inversion is straightforward. Let us here elaborate some on the need for an analytical inversion. The eigenfrequency cannot be identified properly unless we can find an exponentially damped sinusoid, Eq. (3). If we use the full expressions for $\omega_t(\omega)$ and $\beta_t(\omega)$, and use numerical inversion, we may find a motion which looks very much like Eq. (3). However, if there is a weak correspondence, we may loosely argue that the eigenfrequency is time dependent. More about this when we come to an application. Now, it turns out, by trial and error, that by evaluating $\omega_t(\omega)$ and $\beta_t(\omega)$ for the frequency usually named as the adiabatic resonance frequency, ω_a,

$$\omega_a^2 = \frac{P_{b0}}{\rho_0 R_0^2}\left[3\gamma - \frac{2\sigma}{R_0 P_{b0}}\right], \tag{12}$$

we get good estimates of $Q(\omega)$. That is, by comparing the approximation of $Q(\omega)$ with the full expression in terms of the frequency behavior of the magnitude and the phase, a good correspondence is observed. The consequence of this fact is that the inversion can be safely done by means of the simpler rational form of $Q(\omega)$. Furthermore, the procedure presented here can be considered as a new rationale for the approximation of the quantities ω_t and β_t. Now, when $\omega_t \approx \omega_t(\omega_a)$ and $\beta_t \approx \beta_t(\omega_a)$, the inversion yields an expression like Eq. (3), and the result for the eigenfrequency is

$$\omega_{\text{free}}^2 = \frac{\omega_t^2(\omega_a)}{B} - \beta_{\text{free}}^2, \tag{13}$$

where

$$B = 1 + 2R_0[\beta_\nu + \beta_t(\omega_a)]/c_0, \tag{14}$$

and

$$\beta_{\text{free}} = \frac{2[\beta_\nu + \beta_t(\omega_a)] + \omega_t^2(\omega_a)R_0/c_0}{2B}. \tag{15}$$

It is easy to check the above results. For the case of viscous effects only, we have an exact mechanical oscillator analogy with constant parameters, and from the above equations we get: $\beta_{\text{free}} = \beta_\nu$ and $\omega_{\text{free}} = \sqrt{\omega_a^2 - \beta_\nu^2}$, as expected. We obtain an analogous result if we include the approximations discussed above: $\beta_{\text{free}} = \beta_\nu + \beta_t(\omega_a)$ and $\omega_{\text{free}} = \sqrt{\omega_t^2(\omega_a) - \beta_{\text{free}}^2}$. Here, ω_r and β_r are zero.

The adiabatic resonance frequency, ω_a, is frequently used in order to estimate the bubble radius out of observations of free oscillations [Eq. (12) with surface tension neglected].[10,11] Let us as an application of our new results check the accuracy of the above approach. In Fig. 1 the ratio

$\omega_a/\omega_{\text{free}}$ is plotted as a function of equilibrium bubble radius. It is seen that the approximation $\omega_a \approx \omega_{\text{free}}$ is poor for small bubbles and that it is improved for larger bubbles. It is important to note here that our new results are based on an approximation and some caution should be taken. However, it is not just so that the approximated Fourier transform, $Q[\omega, \omega_t(\omega_a), \beta_t(\omega_a)]$, has been found to be applicable, but we have also solved the integral numerically for bubble sizes within the range 0.1 to 1000 μm.[14] The result is that the differences cannot be seen on a reasonably sized graph for bubbles larger than, say, 10 μm. In the range 0.1 to 10 μm, the discrepancy goes from some 10% (overestimate) to essentially zero as the bubble size is increased. Thus, the approximation has been tested in time domain, and the curve shown in Fig. 1 actually underestimates the error obtained while using the adiabatic resonance frequency as an approximation of the transient frequency. Furthermore, the above approach is independent of the model used for the thermal processes. That is, should improved results for $\omega_t(\omega)$ and $\beta_t(\omega)$ occur in the literature, they could in principle be tested within the above method.

Our procedure presented above for finding an expression for the eigenfrequency was formally strict but quite cumbersome, and with this as a solid background, one can address a simpler approach with harmonic motion. By letting the forcing P_e tend to zero in Eq. (4) while keeping ΔR finite, we have that $Q(\omega)$ must tend to infinity. That is, the denominator of $Q(\omega)$ must tend to zero and we have the standard characteristic equation of vibratory systems

$$\omega_c^2 - \omega^2[1 + 2R_0(\beta_\nu + \beta_c)/c_0] + i\omega[2(\beta_\nu + \beta_c)$$
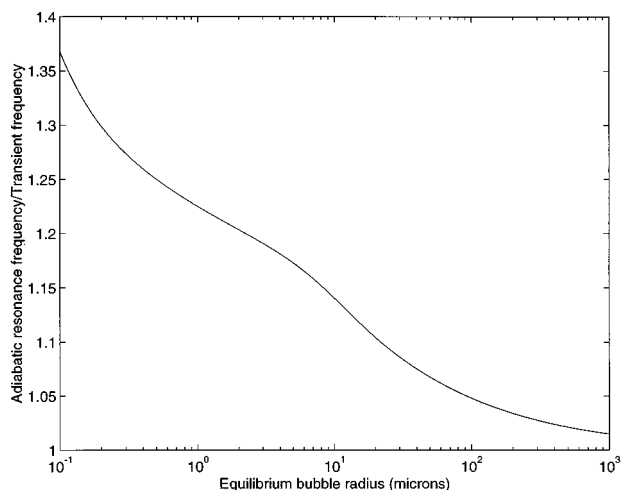$$+ \omega_c^2 R_0/c_0] = 0, \tag{16}$$

FIG. 1. The ratio between the adiabatic resonance frequency and the transient frequency: $\omega_a/\omega_{\text{free}}$, as function of equilibrium bubble radius, R_0.

where $\omega_c = \omega_t(\omega_a)$ and $\beta_c = \beta_t(\omega_a)$. By considering the complex frequency $\omega = \omega_{Re} + i\omega_{Im}$, it is readily shown that the eigenfrequency from the earlier treatment, Eqs. (13) to (15), is recovered. That is, $\omega_{Re} = \omega_{free}$ and $\omega_{Im} = \beta_{free}$.

We conclude this investigation by looking at transient bubble oscillations for the case of liquid-compressibility effects only. Thus, for this case, $\beta_\nu = \beta_t = \sigma = 0$ and $\omega_t = \omega_a$. Here, it is readily found on the basis of the above discussions, Eqs. (13) to (15), that the eigenfrequency is given by

$$\omega_{free}^2 = \omega_a^2 - \beta_{free}^2, \tag{17}$$

and

$$\beta_{free} = \omega_a^2 R_0 / 2 c_0. \tag{18}$$

However, this was obtained analytically without any approximations concerning the compressibility effects. This is interesting since it is sometimes believed that if the quantities ω_b and β_{tot} are frequency dependent, the analogy between a bubble in a liquid and a standard mechanical oscillator is superficial.[5,13] Moreover, these investigators also argue that for the case of frequency-dependent parameters, the bubble behavior is dependent upon the past history of the motion. None of the above statements is thus correct when it comes to the above compressible case. To further visualize this issue, we look at the homogeneous linearized Keller equation[17]

$$\Delta\ddot{r} + \frac{\omega_a^2 R_0}{c_0} \Delta\dot{r} + \omega_a^2 \Delta r = 0. \tag{19}$$

By direct identification, Eqs. (17) and (18) are recovered. The conclusion is that it is not the frequency dependence itself that matters, it is rather the form of it that governs whether or not we have a simple mathematical mechanical oscillator analogy.

III. CONCLUSIONS

The principal results from this work can be given as follows:

(1) By means of an existing model for linear gas bubble pulsation in liquids we have derived an approximate expression for the eigenfrequency. It has been shown that the commonly used adiabatic resonance frequency overestimates the transient frequency of the free bubble oscillations. This is most significant for small bubbles.
(2) A new method for approximating the functions reflecting the thermal processes has been suggested. It makes the approximation of the bubble behavior independent of the frequency contents of any possible excitation. Moreover, the effects due to compressibility of the surrounding liquid are untouched in the above.
(3) It has been demonstrated that for the effects of liquid-compressibility only, the bubble behavior is independent of the past history of the motion. The linearized homogeneous Keller equation represents a full nonsuperficial mechanical oscillator analogy. The existence of frequency dependent parameters for harmonic motion represents a sufficient and not necessary condition for having a superficial oscillator.

[1] E. L. Carstensen and L. L. Foldy, ''Propagation of sound through a liquid containing bubbles,'' J. Acoust. Soc. Am. **19**, 481–501 (1947).
[2] F. E. Fox, S. R. Curley, and G. S. Larson, ''Phase velocity and absorption measurements in water containing air bubbles,'' J. Acoust. Soc. Am. **27**, 534–539 (1955).
[3] L. van Wijngaarden, ''On the equations of motion for mixtures of liquid and gas bubbles,'' J. Fluid Mech. **33**, 465–474 (1968).
[4] D. S. Drumheller and A. Bedford, ''A theory of bubbly liquids,'' J. Acoust. Soc. Am. **66**, 197–208 (1979).
[5] K. W. Commander and A. Prosperetti, ''Linear pressure waves in bubbly liquids: Comparison between theory and experiments,'' J. Acoust. Soc. Am. **85**, 732–746 (1989).
[6] S. Temkin, ''Attenuation and dispersion of sound in bubbly fluids via the Kramers–Kronig relations,'' J. Fluid Mech. **211**, 61–72 (1990).
[7] J. L. Leander, ''Transient wave propagation through bubbly layers via the Foldy–Twersky integral equation,'' J. Acoust. Soc. Am. **95**, 2378–2386 (1994).
[8] R. Omta, ''Oscillations of a cloud of bubbles of small and not so small amplitude,'' J. Acoust. Soc. Am. **82**, 1018–1033 (1987).
[9] L. d'Agostino and C. E. Brennen, ''Acoustical absorption and scattering cross sections of spherical bubble clouds,'' J. Acoust. Soc. Am. **84**, 2126–2134 (1988).
[10] H. C. Pumphrey and L. A. Crum, ''Free oscillations of near-surface bubbles as a source of the underwater noise of rain,'' J. Acoust. Soc. Am. **87**, 142–148 (1990).
[11] J. S. Stroud and P. L. Marston, ''Optical detection of transient bubble oscillations associated with the underwater noise of rain,'' J. Acoust. Soc. Am. **94**, 2788–2792 (1993).
[12] J. A. Nystuen and H. Medwin, ''Underwater sound produced by rainfall: Secondary splashes of aerosols,'' J. Acoust. Soc. Am. **97**, 1606–1613 (1995).
[13] A. Prosperetti, L. A. Crum, and K. W. Commander, ''Nonlinear bubble dynamics,'' J. Acoust. Soc. Am. **83**, 502–514 (1988).
[14] J. L. Leander, ''Transient scattering from small bubbles,'' in *Proceedings of Ultrasonics World Congress 1995* (Berlin, Germany), pp. 615–618.
[15] P. M. Morse and K. U. Ingaard, *Theoretical Acoustics* (Wiley, New York, 1981).
[16] J. L. Leander, ''Transient vibrations of linear oscillators via Fourier integrals and complex analysis,'' Report TRITA-TAK-9444 (1994).
[17] J. B. Keller and I. I. Kolodner, ''Damping of underwater explosion bubble oscillations,'' J. Appl. Phys. **27**, 1152–1155 (1952).

TV holography measurements of underwater acoustic fields

Rolf Rustad, Ole Johan Løkberg, Hans Magne Pedersen, Kristin Klepsvik, and Trude Støren

Applied Optics Group, Institute of Physics, Norwegian University of Science and Technology, 7034 Trondheim, Norway

Members of our group have previously demonstrated that TV holography may be used to measure the amplitude and the phase distributions of a sound field in air. This technique is not limited to acoustic fields in air, but may be applied to acoustic fields in any transparent medium. Its applicability to ultrasonic fields in water is demonstrated in this Letter. © *1997 Acoustical Society of America.* [S0001-4966(97)01309-X]

PACS numbers: 43.58.Vb, 43.20.Ye, 43.35.Sx [SLE]

INTRODUCTION

Ultrasonic fields in water are investigated through a wide variety of methods, like scanning hydrophones or pulse echo techniques. Optical investigation offers the advantage of noninvasive measurements. Methods based on Raman–Nath diffraction[1] are well known, but their use is limited regarding the intensity and wavelength of the acoustic fields under investigation. TV holography measurements[2,3] do not rely on diffraction and are applicable to acoustic fields over a wide range of frequencies, from some 100 Hz into the MHz range. The lower limit in acoustic intensity is not frequency dependent, but is determined by the general signal to noise ratio of the system. Pressure amplitudes less than 100 Pa are measurable under certain conditions. Furthermore, there are no limitations to the shape of the acoustic waves. Finally, the experimental procedure is fairly simple and robust, so that one may imagine using this technique outside the laboratory.

We demonstrate this technique through measurements of the continuous acoustic field produced by a circular transducer submerged in fresh water in a tank.

I. THEORY

A light ray of wavelength λ traversing a uniform acoustic field of width L, angular frequency Ω, and phase Φ is subject to a periodic variation of its phase ϕ equal to

$$\phi(t) = \frac{2\pi\mu L}{\lambda}\cos(\Omega t + \Phi) = \nu\cos(\Omega t + \Phi), \qquad (1)$$

where ν is the familiar Raman–Nath parameter and μ is the maximum variation of the refractive index. This phase variation may be measured quantitatively by time averaged TV holography.[2] The result is equal to that produced by a surface vibrating with amplitude $a_0 = \mu L$. Knowing μ, L, and the piezo-optic constant $p_{op} = \partial n / \partial P$, the pressure amplitude may be calculated.

The detection limit for time averaged TV holography with our current configuration is on the order of $a_{0,min} \approx 1$ nm. Using[4] $p_{op} = 1.51 \times 10^{-10}$ Pa^{-1} gives $\Delta p_{min} \approx 6$ Pa m/ L, where L is the length of the path followed by a light ray through the field.

II. EXPERIMENT

A diagram of the experimental setup is shown in Fig. 1. For our underwater experiment we used a tank equipped with four windows, one on each side. Two windows are needed for the measurements, and the other two are used for inspection and alignment. The tank has a volume of approximately 0.5 m^3 and was filled with tap water at room temperature. A Retra 1000 TV holography unit with a 10-mW HeNe laser at wavelength 632.8 mm was placed in front of one of the windows. An electro-optical modulator in the reference arm of the interferometer allowed phase modulation of the reference beam. A detailed description of the Retra 1000 is given in Ref. 3.

The measurement beam is directed out of the interferometer and projected into the tank. It traverses the water with the acoustic field and illuminates a retroreflector outside the opposing window. Retroreflection eliminates cumbersome alignment procedures needed when using mirrors. The resulting speckle field traverses the tank a second time as it is imaged onto the CCD camera of the interferometer. Each pixel of the camera thus measures the accumulated change in the optical phase of a ray that has traversed the acoustic field twice. We assume the variations in refractive index to be
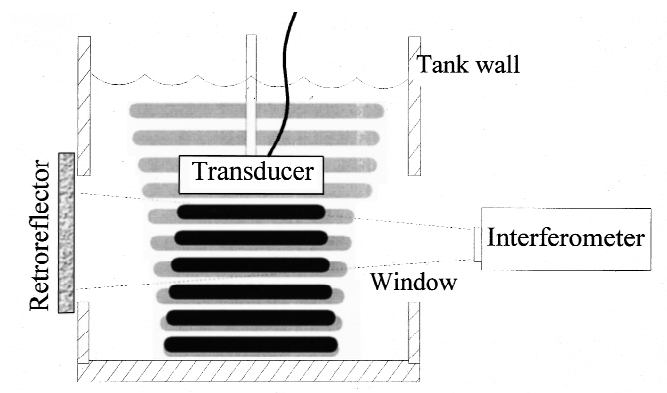

FIG. 1. A cross section through the center of our experimental setup. A transducer is submerged in fresh water in a tank with windows on all sides. The transducer emits a nearly planar wave which is reflected from the bottom of the tank. Projections of this acoustic field are recorded by the interferometer outside the window to the right. A retroreflector is placed outside the window to the left.
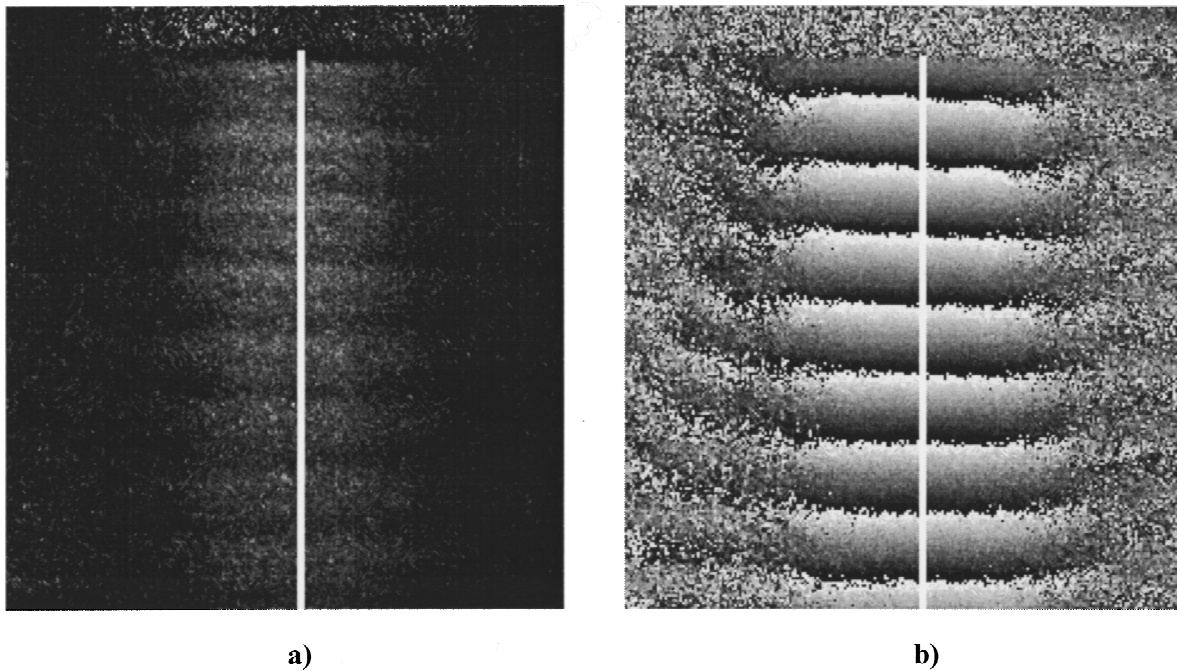
FIG. 2. A projection of the amplitude (a) and the phase (b) of an acoustic field in water produced by a transducer exited at 141.5 kHz, viewed from the side. The transducer is seen at the top of the images and is radiating downwards. The amplitude image (a) is a gray scale representation where lighter gray indicates higher acoustic intensity. In the phase image (b), black and white corresponds to −180 and 180 degrees, respectively. The phase is wrapped. Traces of the amplitude and the phase along the indicated line is presented in Fig. 3.

small enough for the light to propagate undiffracted and along straight lines. The inline illumination and observation of the interferometer ensures that each ray follows the same path through the acoustic field both times.

A circular hydroacoustic transducer with a diameter of 69 mm and a planar emitting face was submerged into the water and operated in continuous mode. The emitting face was directed toward the bottom of the tank, which was approximately 50 cm away. At 141.5 kHz the transducer emits a nearly planar wave which is reflected from the bottom of the tank.

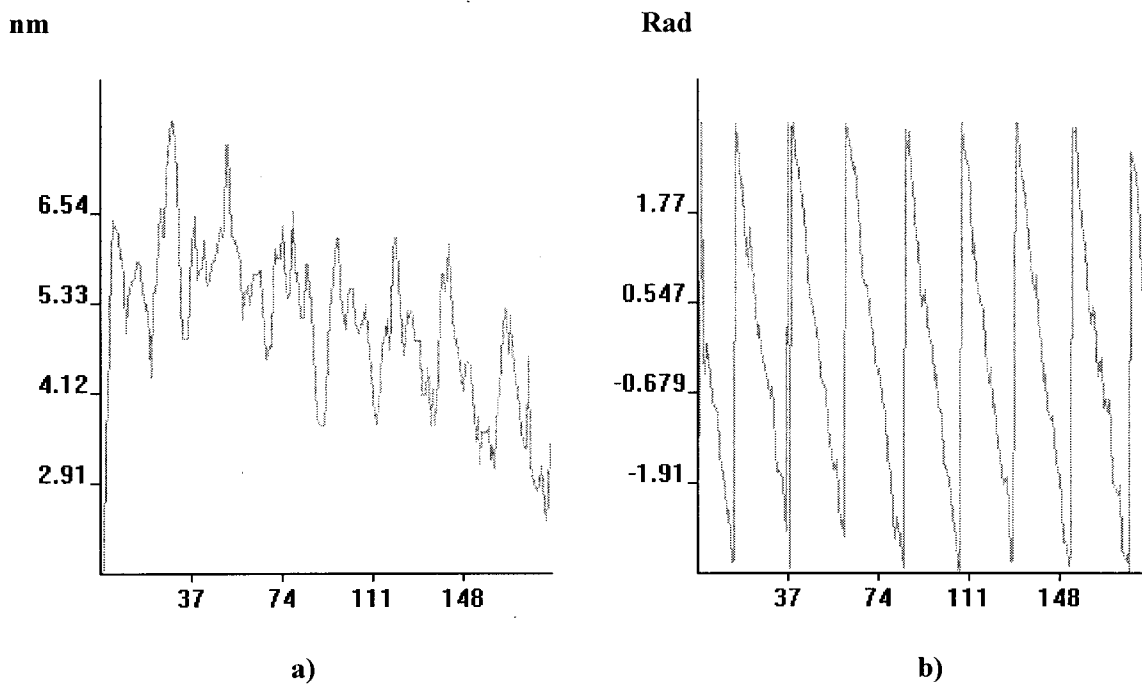Figure 2 shows the recorded projections of the ampli-

FIG. 3. Traces of the amplitude (a) and the phase (b), of the projection of the acoustic field in Fig. 2 along the lines indicated in the figure. The amplitude is in units of equivalent surface vibration in nm. The x ordinate is in pixel units, with one pixel being equivalent to approximately 0.44 mm.

tude and the phase of this acoustic field viewed from the side. Part of the transducer is visible at the top of each image as a rectangle of salt and pepper noise. Figure 2(a) is a gray scale representation of the measured projection of the sound intensity. Light gray corresponds to highest intensity. The gray scale is manipulated to enhance visibility. Figure 2(b) is a gray scale representation of the measured projection of the phase. Black and white corresponds to -180 and 180 degrees, respectively. The phase is wrapped. Both images present unfiltered data.

Figure 3 shows traces of the amplitude and the phase along the line indicated in Fig. 2. The line is parallel to the direction of propagation of the acoustic field. The amplitude, Fig. 3(a), is presented in units of equivalent surface vibration, in nm. The x ordinate is in pixel units, with each pixel being approximately 0.44 mm wide.

The apparent modulation of the intensity is caused by integration through regions of opposing phase at the edge of the acoustic field.[2] The wave may also have been distorted by interference with its own reflection caused by the bottom of the tank.

III. CONCLUSION

We have shown that projections of the amplitude and the phase of an acoustic field in water may be measured using TV holography. Similar measurements of sound fields in air have been used for tomographic reconstructions to yield a full three-dimensional mapping of the amplitude and the phase of the field.[3] Our group is currently working on the application of these tomographic techniques to acoustic fields in water and proper calibration routines to yield reliable absolute values for both amplitude and phase at any point in the field.

[1] C. V. Raman and N. S. N. Nath, ''The diffraction of light by high frequency sound waves: Part I–II,'' Proc. Ind. Acad. Sci. IIA, 406–420 (1936).
[2] O. J. Løkberg, ''Sound in flight: Measurement of sound fields by use of TV holography,'' Appl. Opt. **33**, 2574–2584 (1994).
[3] M. Espeland, O. J. Løkberg, and R. Rustad, ''Full field tomographic reconstruction of sound fields using TV holography,'' J. Acoust. Soc. Am. **98**, 280–287 (1995).
[4] W. A. Riley and W. R. Klein, ''Piezo-optic coefficients of liquids,'' J. Acoust. Soc. Am. **42**, 1258–1261 (1967).

The effect of a free-field auditory target's motion on its detectability in the horizontal plane[a]

Xin Xiao and D. Wesley Grantham[b]
Department of Hearing and Speech Sciences, Vanderbilt Bill Wilkenson Center for Otolaryngology and Communication Sciences, Vanderbilt University Medical Center, 1114 19th Avenue South, Nashville, Tennessee 37212

Masked thresholds were measured for five subjects in an anechoic chamber. The signal (a 500-Hz tone, an 8000-Hz tone, or a broadband noise) was presented from a movable loudspeaker located in the horizontal plane at ear level. A continuous noise masker came from two fixed loudspeakers at $\pm 14°$ azimuth. In different threshold runs the signal was either stationary (positioned at 0° or at 15° azimuth), or it moved during its on-time (at a velocity of 30°/s or 90°/s) through an arc from $-15°$ to $+15°$. There was no evidence that any of the moving signals was more detectable than a stationary signal when threshold variation due to position changes of the stationary signal was taken into account. Likewise, there was no evidence that signal velocity affected detectability when signal duration was taken into account. These results agree with previous reports obtained with simulated moving signals presented via earphones and suggest that if horizontal motion-sensitive mechanisms exist in the auditory system, such mechanisms do not contribute to signal detectability. © 1997 Acoustical Society of America. [S0001-4966(97)04209-4]

PACS numbers: 43.66.Dc, 43.66.Mk, 43.66.Pn [RHD]

INTRODUCTION

Detectability of auditory signals is clearly critical for most organisms, whose daily survival depends on the ability to find and identify prey, predator, offspring, and mates. Since many of the critical sounds in an organism's environment are produced by objects or creatures that are in motion, an important question is whether moving sounds are more detectable by mammals (in particular, humans) than stationary sounds.

In the case of vision, it was observed early in the 19th century that a moving visual target is more detectable than a stationary one (Boring, 1942). More recent experiments have shown that the detectability of a visual moving object is a complex function of the object's velocity, spatial frequency composition, and the portion of the retina that is stimulated. Under some circumstances we can know that something is moving without knowing what it is (Sekuler *et al.*, 1978).

Does the motion of a sound source similarly facilitate its detectability? Results from studies conducted with earphones suggest that moving auditory signals do *not* elicit a lower masked threshold than do stationary signals if account is taken of the variation in threshold than can occur when a stationary signal is presented from different locations in space (Wilcott and Gales, 1954; Robinson, 1971; Grantham and Luethke, 1988). However, all of these studies *simulated* motion by varying the interaural time differences (ITDs) of signals presented to subjects over earphones; this technique thus does not incorporate the changing interaural level differences (ILDs) or changing spectral cues that are also known to be important for localizing sounds in free field (Grantham, 1995).

Although there is thus no evidence that simulated motion enhances the detectability of signals in human beings, physiological research has suggested that some animals are more sensitive to moving than to stationary auditory signals. Thus, Takahashi and Keller (1992) reported that responses of single neurons in the barn owl's inferior colliculus (IC) to the simulated azimuth of auditory targets was generally more resistant to masking noise when the targets were moving than when they were stationary. Likewise, Spitzer and Semple (1991) found that the response rate of IC neurons in gerbil and cat to interaural phase differences (IPDs) was consistently higher when the IPD was varying than when it was static.[1] In behavioral studies, both Simmons (1981) and McCarty and Jen (1983) have reported that echo-locating bats had lower masked thresholds for echoes when those echoes were moving (in depth) than when stationary. Thus, at least in some species and under some circumstances, there is evidence that moving signals are more detectable than stationary signals.

We reasoned that if horizontal-plane motion can enhance the detectability of an auditory signal for human listeners, the most favorable condition for discovering and investigating the advantage should be in a realistic environment, where all cues for spatial processing (ITDs, ILDs, and spectral cues) are preserved. In the present experiment we investigated the effect of motion on the detectability of various auditory signals in the free field. Two questions were addressed: (1) Is a moving auditory target more detectable than a comparable stationary target, when the threshold changes due to a stationary target's different positions are taken into account? (2) For moving targets that traverse a given constant angular extent at different velocities, is there an effect of velocity on target detectability?
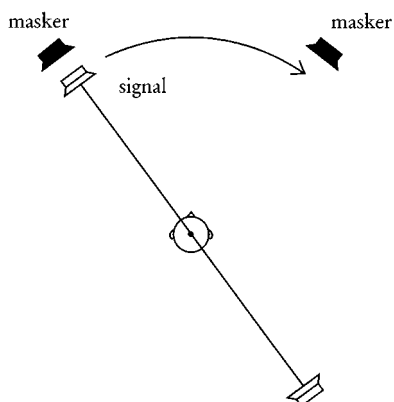
FIG. 1. Loudspeaker configuration for signal detection task. Continuous noise masker was presented from the two stationary loudspeakers at ±14° azimuth (filled icons). Signal was presented from one of the movable loudspeakers (open icons).

I. METHOD

A. Subjects

Five adult subjects, one female and four male, aged 20–30, with little or no previous experience in psychoacoustic experiments, participated in this experiment. Subjects had clinically normal hearing ability bilaterally.

B. Testing environment and stimuli

The experiment was conducted in a darkened anechoic chamber (6 m×6 m×6 m). Subjects were run individually and were instructed to sit upright in the chair in the center of the chamber and face forward without tilting the head to the side or forward during the run. Although the subjects were not monitored for head movement, they were reminded prior to each run to sit still and maintain the head in a forward orientation. Four loudspeakers (JBL-8110H) were employed (Fig. 1): Two stationary loudspeakers were positioned 1.7 m in front of the subject, 0.85 m above the floor at −14° and +14° azimuth, respectively. These two stationary loudspeakers delivered the masking noise. Two other loudspeakers were suspended from opposite ends of an overhead boom, which could rotate around a pivot point over the subject's head. These two (signal) loudspeakers could trace a horizontal circular path around the subject at a distance of 1.48 m at ear level (1.10 m above the floor). Thus, the two movable loudspeakers were slightly closer to the subject and slightly higher than the two stationary loudspeakers.

The masking noise consisted of two independent broadband (100–22 000 Hz) Gaussian noises (one to each masker loudspeaker), presented continuously at a level of 40 dB SPL. In preliminary informal testing, masker level appeared to have no consistent effect on the pattern of thresholds; here, a relatively low level (40 dB SPL) was chosen in order to simulate a "quiet" environment, yet be of sufficient intensity to mask uncontrolled low-level ambient noise generated by the room's air conditioning system. Uncorrelated noises were employed at the two loudspeakers in order to provide a background noise with some spatial width (Morimoto *et al.*, 1993), thus simulating the conditions that might exist in a natural environment.

Three different signals were employed: a 500-Hz tone, a 8000-Hz tone, and a broadband (100–10 000 Hz) noise.

C. Procedure

For each signal, masked threshold was determined in five conditions:

(1) moving at a velocity of 30°/s for a duration of 1000 ms (thus, traversing an arc of 30°, from −15° azimuth to +15° azimuth);
(2) stationary at 0° azimuth, for a duration of 1000 ms;
(3) moving at a velocity of 90°/s for a duration of 333.3 ms (thus, traversing an arc of 30°, from −15° to +15° azimuth);
(4) stationary at 0°, for a duration of 333.3 ms;
(5) stationary at 15° azimuth, for a duration of 333.3 ms.

Thus, the moving signals (conditions 1 and 3) always traversed an arc that extended approximately from the leftmost masker position to the rightmost masker position. The 30° arc traversed by the moving signals was well above the threshold angle for motion detectability, so that motion was always clearly perceived when these signals were presented at audible levels [minimum audible movement angles for these types of signals presented at the velocities employed here range from 4°–20° (Perrott and Tucker, 1988; Chandler and Grantham, 1992)]. Stationary signals were presented either from midline (conditions 2 and 4) or from a position near the rightmost masker (condition 5). The order of presentation of the five conditions and the three signal types was counterbalanced across subjects.

Masked thresholds were determined in an adaptive, single-interval, two-alternative forced-choice task: prior to each threshold run, the boom was either positioned so that one of the signal loudspeakers was at the desired stationary angle (0° or 15°), or the boom was accelerated to the desired velocity (90°/s or 30°/s). The two-component masker remained on continuously through a run. On half the trials (randomly determined) the signal was pulsed on during the observation interval which was marked by a LED on the subject's response box; on the other half of the trials the signal was not presented. The subject had to respond on each trial "yes" or "no" by pressing a button to indicate whether he or she heard the signal. Immediate feedback was provided. In the case of the moving signal, the signal (on signal trials) was presented from one of the two moving loudspeakers (i.e., whichever one came into position first); for the stationary signals, only one of the loudspeakers was employed throughout the run.

Each run began with the signal level well above threshold, so that the subject could easily detect it. After each three successive correct responses, the signal was attenuated, and after each incorrect response the signal level was increased, thus tracking the 79% correct level (Levitt, 1971). Step size initially was 5 dB, and changed to 2 dB after the first two reversals of signal level. The run terminated after ten reversals in signal level, and threshold was estimated as the average signal level from the last eight reversal points. Runs for which the standard deviation computed across reversal points was greater than 4.0 dB were discarded.
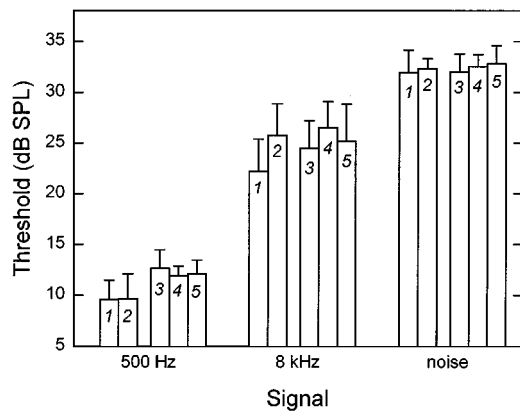
FIG. 2. Average masked thresholds across the five subjects for the three signals (shown on the abscissa) and the five conditions investigated. Error bars indicate one standard deviation. The numbers on the bars correspond to the five conditions listed in the text.

Subjects were given at least one session of practice prior to data collection. Twelve to sixteen runs were completed in a 1.5-h session, in which frequent breaks (every 15–20 min) were allowed. Final thresholds were taken as the medians of at least four runs in each condition (not all four runs were completed on the same day).

II. RESULTS AND DISCUSSION

Average thresholds across the five subjects are shown in Fig. 2 for the three signal types (indicated across the abscissa) and the five conditions of study (shown as different bars for each signal type). Overall masked thresholds, computed across subjects and conditions, were 11.2 dB SPL for the 500-Hz tone, 24.8 dB SPL for the 8-kHz tone, and 32.3 dB SPL for the noise signal. The primary concern of the present study is the *differences* between the thresholds for selected pairs of conditions, which we address in the following sections.

A. Detectability of moving versus stationary targets

Threshold differences for detecting moving versus stationary targets are plotted in Fig. 3 for the three signal types (indicated along the abscissa). The data points show the mean difference scores across the five subjects; error bars show the standard deviation across subjects. On this axis system, positive difference scores indicate an advantage for detecting the moving signal.

The three different symbols in the figure denote three different comparisons that were made among the five conditions employed: the circles show the difference in detectability of the slow signal (30°/s) and the stationary target at 0° having the same duration (1000 ms: condition 1 versus 2); the squares show the difference in detectability of the fast signal and the stationary target at 0° having the same duration (333 ms: condition 3 versus 4); and the triangles show the difference in detectability of the fast signal and the stationary target at 15° having the same duration (333 ms: condition 3 versus 5). The figure indicates that for the 500-Hz and noise signals there were no threshold differences between moving and stationary targets for any of the comparisons.
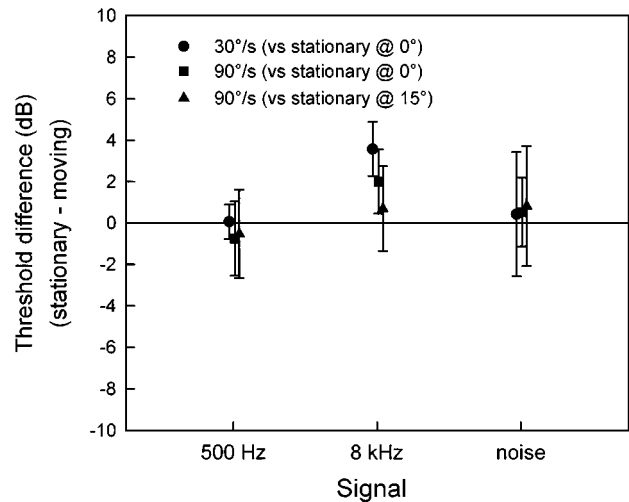

FIG. 3. Difference in threshold between stationary and moving signals (positive difference indicates advantage for moving signal). Means and standard deviations across the five subjects are plotted for the three signal types indicated along the abscissa.

The results for the 8000-Hz signal were more interesting. Here the 30°/s target was more detectable than the stationary target at midline by an average of 3.6 dB [$F(1,4)$ $=37.33$; $p<0.01$]; likewise, the 90°/s target was more detectable than the midline stationary target by an average of 2.0 dB [$F(1,4)=8.39$; $p<0.01$]. At first glance, these results appear to suggest that the 8000-Hz moving signal *was* more detectable than a stationary signal. However, when the comparison was made between the moving target and a stationary target positioned at $+15°$ (triangles), there was no longer any threshold advantage for the moving signal [$F(1,4)$ $=0.57$; $p=0.49$]. Evidently, the moving targets were more detectable than stationary targets at 0° azimuth, not because they were in motion *per se*, but because the moving signal passed through an azimuth ($+15°$) that naturally provided a better signal-to-noise ratio than that obtained at 0° (e.g., see Saberi *et al.*, 1991).

B. Effect of velocity on signal detectability

Conditions 1 and 3 both involve a moving target that traverses an arc from $-15°$ to $+15°$ azimuth, but at different velocities (and thus, for different durations). A comparison of performance in these two conditions can potentially reveal whether signal velocity has an effect on detectability. Figure 4 plots the mean threshold differences across subjects between the these two conditions (shown as circles: positive differences indicate an advantage for detecting the slower target). There is no apparent difference for the 8000-Hz or for the noise signal, but for the 500-Hz signal the slower target is more detectable than the faster by 3.1 dB [$F(1,4)$ $=47.52$; $p<0.01$].

This difference might be attributable to the different velocities, the different durations, or both. An independent assessment of the effect of duration may be obtained by comparing detectability of the two *stationary* targets that have these durations (conditions 2 and 4). This comparison is shown in Fig. 4 as the squares; for the 500-Hz signal the 1000-ms target is more detectable than the 333-ms target by
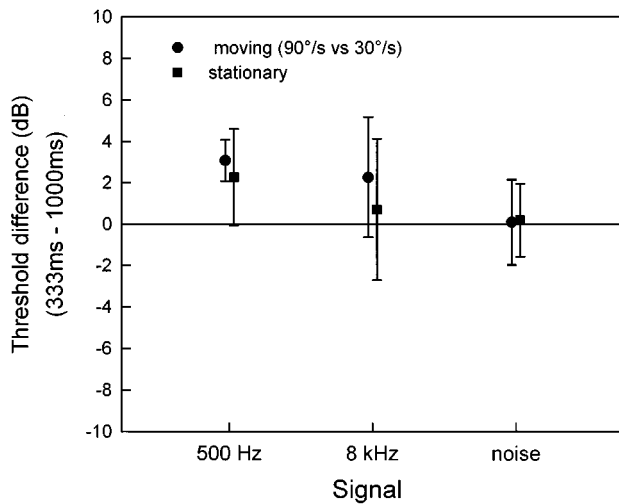
FIG. 4. Difference in threshold between a long (1000 ms) and a short (333 ms) signal (positive difference indicates advantage for the longer signal). Means and standard deviations across the five subjects are plotted for the three signal types indicated along the abscissa.

2.3 dB $[F(1,4) = 7.89; p < 0.05]$. The fact that the advantage of the longer-duration signal (with stationary targets) is almost as great as the advantage of the slower moving target suggests that the latter effect is not based on velocity, but rather on the different durations of the signals.

III. CONCLUSIONS

For the conditions investigated there is no evidence that motion of auditory targets in the free-field horizontal plane renders them any more (or less) detectable than stationary targets. Although there are cases in which a moving target's threshold is lower than that of a stationary target, the advantage is not due to motion of the target *per se*, but to the fact that the moving target may pass through azimuths that yield a better signal-to-noise ratio than that of the stationary target. There is also no evidence that detectability of targets that traverse a given arc in the horizontal plane is affected by the velocity of the target, once signal duration has been taken into account.

These results of free-field signal detectability support previous reports obtained with earphones that dynamic targets are no more detectable than stationary targets. They are also consistent with findings that the spatial discrimination of horizontally moving sounds can be almost completely accounted for by a static comparison of the moving targets' endpoints (Grantham, 1986, 1997; however, see also Perrott and Marlborough, 1989). Thus, although physiological evidence suggests the existence of motion-sensitive mechanisms in the auditory systems of some animals, human psychophysical data have not consistently identified a strong behavioral role for such mechanisms.

ACKNOWLEDGMENTS

[1]It should be noted that not all physiological studies have found a threshold advantage for moving targets. For example, Wagner *et al.* (1994) found no differences in response thresholds to moving versus stationary targets in motion-sensitive neurons in the owl's midbrain.

Boring, E. G. (**1942**). *Sensation and Perception in the History of Experimental Psychology* (Appleton, New York), p. 594.

Chandler, D. W., and Grantham, D. W. (**1992**). ''Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity,'' J. Acoust. Soc. Am. **91**, 1624–1636.

Grantham, D. W. (**1986**). ''Detection and discrimination of simulated motion of auditory targets in the horizontal plane,'' J. Acoust. Soc. Am. **79**, 1939–1949.

Grantham, D. W. (**1995**). ''Spatial hearing and related phenomena,'' in *Handbook of Perception and Cognition: Hearing*, edited by Brian C. J. Moore (Academic, San Diego).

Grantham, D. W. (**1997**). ''Auditory motion perception: Snapshots revisited,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by T. R. Anderson and R. H. Gilkey (Erlbaum, Mahwah, NJ), pp. 295–313.

Grantham, D. W., and Luethke, L. E. (**1988**). ''Detectability of tonal signals with changing interaural phase differences in noise,'' J. Acoust. Soc. Am. **83**, 1117–1123.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

McCarty, J. K., and Jen, P. H. S. (**1983**). ''Bats reject clutter interference for moving targets more successfully than for stationary ones,'' J. Comp. Physiol. **152**, 447–454.

Morimoto, M., Iida, K., and Furue, Y. (**1993**). ''Relation between auditory source width in various sound fields and degree of interaural cross-correlation,'' Appl. Acoust. **38**, 291–301.

Perrott, D. R., and Marlborough, K. (**1989**). ''Minimum audible movement angle: Marking the end points of the path traveled by a moving sound source,'' J. Acoust. Soc. Am. **85**, 1773–1775.

Perrott, D. R., and Tucker, J. (**1988**). ''Minimum audible movement angle as a function of signal frequency and the velocity of the source,'' J. Acoust. Soc. Am. **83**, 1522–1527.

Robinson, D. E. (**1971**). ''The effect of interaural signal-frequency disparity on signal detectability,'' J. Acoust. Soc. Am. **50**, 568–571.

Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., and Perrott, D. R. (**1991**). ''Free-field release from masking,'' J. Acoust. Soc. Am. **90**, 1355–1370.

Sekuler, R., Pantle, A., and Levison, E. (**1978**). ''Physiological basis of motion perception,'' in *Handbook of Sensory Physiology, Vol. VIII Perception*, edited by R. Held, H. W. Leibowitz, and H.-L. Teuber (Springer-Verlag, Berlin).

Simmons, J. A. (**1981**). ''Detection of range motion in bat sonar,'' J. Acoust. Soc. Am. Suppl. 1 **69**, S112.

Spitzer, M. W., and Semple, M. N. (**1991**). ''Interaural phase coding in auditory midbrain: Influence of dynamic stimulus features,'' Science **254**, 721–724.

Takahashi, T. T., and Keller, C. H. (**1994**). ''Representation of multiple sound sources in the owl's auditory space map,'' J. Neurosci. **14**, 4780–4793.

Wagner, H., Trinath, T., and Kautz, D. (**1994**). ''Influence of stimulus level on acoustic motion-direction sensitivity in barn owl midbrain neurons,'' J. Neurophysiol. **71**, 1907–1916.

Wilcott, R. C., and Gales, R. S. (**1954**). ''Comparison of the masked thresholds of a simulated moving and stationary auditory signal,'' J. Exp. Psychol. **47**, 451–456.

Xiao, X., and Grantham, D. W. (**1996**). ''The effect of a free-field auditory target's horizontal motion on its detectability,'' J. Acoust. Soc. Am. **99**, 2597(A).