CODEN: JASMAN

ISSN: 0001-4966

The Journal of the Acoustical Society of America

Vol. 115, No. 2 February 2004

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 115, NO. 2, FEBRUARY 2004

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 115, NO. 2, FEBRUARY 2004

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 115, NO. 2, FEBRUARY 2004

CONTENTS—*Continued from preceding page*

ACOUSTICAL NEWS—USA

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

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

Preliminary Notice: 75th Anniversary (147th) Meeting of the Acoustical Society of America

Twenty-three papers were presented and 183 people attended the first meeting of the Acoustical Society on 10–11 May 1929 in New York City. The 147th meeting—75 years later—will be held at the Sheraton New York Hotel and Towers on 7th Avenue, Monday through Friday, 24–28 May 2004. A block of rooms has been reserved at the Sheraton. In addition to an outstanding technical program, there will be a celebration with the theme, ''Glorious Past—Looking Forward.''

The celebration will begin on Tuesday night with a 75th ''celebratory banquet'' which will feature a video with recollections by past presidents covering highlights from the Society's 75 years. This will be followed by a number of events on Wednesday. In the morning there will be a ''celebratory look'' to the future of the Society at City Center followed by a reception for Fellows of the Society. A number of special tours have been scheduled for the afternoon, and in the evening there will be an organ concert at St. Thomas Church.

The technical program will include over 40 special sessions, two distinguished lectures, and a tutorial by Leo Beranek, all of which are described in this Meeting Announcement. For further up-to-date information about the technical program, please refer to the ''Meetings Section'' on the ASA Home page at ⟨http://asa.aip.org/meetings.html⟩. Also, you may register for the meeting online at ⟨http://asa.aip.org⟩.

We hope you will be able to join us in this wonderful technical program and celebration in the exciting city where the Society began 75 years ago.

Ilene J. Busch-Vishniac, President
Charles E. Schmid, Executive Director

Technical Program

The technical program will consist of lecture and poster sessions. Technical sessions will be scheduled Monday through Friday, 24–28 May 2004, but there will be no technical sessions on Wednesday, 26 May, due to the all-day celebration event. The special sessions described below will be organized by the ASA Technical Committees.

Special Sessions

Acoustical Oceanography (AO)
D. Van Holliday special session on acoustical measurements of marine organisms (Joint with Animal Bioacoustics)
In honor of Van Holliday's contributions to the use of acoustics in the study of marine life at multiple trophic levels and for his leadership in fisheries acoustics
Ocean acoustics of earthquakes
Focus on the characteristics, physical mechanisms, and applications of ocean acoustic observations of earthquakes

Animal Bioacoustics (AB)
Infrasonic communication by animals: Signal propagation, generation, reception, and function
Investigate the propagation of infrasonic signals, covering the generation and reception of infrasonic signals by various animal taxa
Natural acoustic behavior of animals: Session in memory of Donald R. Griffin
Natural use of sound by animals, including sensing the environment, orientation, and communication

Architectural Acoustics (AA)
Acoustics of virtual environments
What makes a virtual environment convincing—includes teleconferencing and video games and spatially distributed performance environments
Alternative acoustic environments for performing arts presentations
Present alternative acoustic environments for performing artists
Bell Labs and acoustics (Joint with Engineering Acoustics)
Review of Bell Labs and acoustics with regard to the Society
Effect of room acoustic environment on human productivity/performance (Joint with Noise)
How do room acoustic conditions affect human productivity/performance in various environments?
Methods to quantify opera house acoustics
Test methods for opera house acoustics; the important acoustical factors and measurements of opera house acoustics
Multisensory integration and the concert experience: How visual input affects what we hear
How visual stimuli are involved in how we perceive the acoustical experience in concert halls
Theatres for drama performance—Another two decades (1984–2004)
Poster session—similar to the concert halls poster session in 2002 with the potential to create a publication

Biomedical Ultrasound/Bioresponse to Vibration (BB)
Elasticity imaging
Techniques to image structures in the body based on variations in elastic properties such as shear modulus
High frequency imaging (Joint with Signal Processing in Acoustics)
Use of high frequency ultrasound to image biological structures
High intensity focused ultrasound (Joint with Physical Acoustics)
Use of high intensity focused ultrasound (HIFU) for therapeutic applications in the body

Education in Acoustics (ED)
Apparatus for teaching acoustics: 1929 and before
Live demonstrations of photographic illustrations of apparatus used to teach acoustics in 1929 and before
Careers in acoustics
Introduce acoustics careers to high school guidance counselors—local educators invited
Hands-on demonstrations for high school students
20 hands-on demonstrations for 75 high school students
Student poster session
Open to students including those from local high schools

Engineering Acoustics (EA)
Acoustic payloads of underwater vehicle acoustics
(Joint with Underwater Acoustics)
Development of acoustic payloads for unmanned undersea vehicles

Musical Acoustics (MU)
Digital signal processing methods for restoring, enhancing, and manipulation of music recordings (Joint with Signal Processing in Acoustics)
Topics include noise reduction of recordings, improving fidelity and ambience and manipulation of music recordings
Musical instruments developed during the ASA era
Acoustics of musical instruments developed after the formation of ASA in 1929
Neurophysiology of playing a musical instrument
(Joint with Psychological and Physiological Acoustics)
Focus on speed, accuracy, and strength of the reflex–motor system to control musical instruments, including the voice

 0001-4966/2004/115(2)/431/8/$20.00

New research on pre-1929 instruments
Research on instruments that fell out of use before the ASA era

Noise (NS)

On the occasion of his 90th birthday, special session to honor the contributions of Leo L. Beranek to acoustics and teaching
(Cosponsored by all Technical Committees and ASA Committees on Education in Acoustics and Standards)
Historical reviews of contributions by Leo Beranek to several technical areas, and to standards and teaching
Noise and society
(Joint with Psychological and Physiological Acoustics)
Review the adverse impacts of noise on hearing, mental and physical health, children's development, and quality of life and explore the educational and advocacy programs that are addressing these issues
Noise control enforcement in New York City
(Joint with ASA Committee on Standards)
A look at approaches to noise enforcement and their effectiveness
Noise impact evaluation—Old and new
(Joint with Psychological and Physiological Acoustics)
Analyzing noise effects; soundscape and/versus community noise measurements
Noise in large cities
Trends in noise levels in urban areas, noise sources, singular noise problems in specific cities, mitigation measures

Physical Acoustics (PA)

Recent advances in buried landmine detection
Recent technological and scientific advances in buried landmine detection
Robert E. Apfel memorial session
(Joint with Biomedical Ultrasound/Bioresponse to Vibration)
A celebration of the life and contributions of Robert Apfel to the field of physics

Psychological and Physiological Acoustics (PP)

Compression in hearing
Focus on compression observed in mammalian cochlea; perceptual effects and impact of hearing loss
Perceptual organization of sound
Overview of psychophysical and physiological aspects of perceptual organization of sound sources
The perception of complex sounds: Honoring the contributions of Charles S. Watson (Joint with Speech Communication)
Invited and contributed papers to review current research regarding the perception of complex sounds

Signal Processing in Acoustics (SP)

Advances in sonar and imaging techniques including interferometric, synthetic, and tomographic apertures (Joint with Underwater Acoustics)
Presentations on advances in sonar and acoustic imaging techniques and discussion of pacing technology issues
Inverse problems in seismic signal processing
(Joint with Acoustical Oceanography)
Survey of current academic and industrial methods and applications for solving the seismic inversion problem
Underwater acoustic communications (Joint with Underwater Acoustics)
New techniques for improving performance and reducing environmental degradation through adaptive modulation, demodulation channel estimation, networks, simulation tools, and performance analysis

Speech Communication (SC)

Forty Years of Voice Onset Time (VOT)
Invited and contributed papers on the use of Voice Onset Time to commemorate the first publication on VOT by Leigh Lisker and Arthur Abramson
Speech perception and production in monolingual and bilingual acquisition
Invited and contributed papers on bilingual acquisition to explore research on features in perception and production skills of those who acquire a second language in childhood

ASA Committee on Standards (ASACOS)

Role of standards in ASA
Introductory presentations followed by presentations from Technical Committees on role of standards to those Technical Committees

Structural Acoustics and Vibration (SA)

Experimental modal analysis
Survey of concepts and algorithms for extrapolations of vibration properties from measurements in the time and frequency domain
Urban transit noise (Joint with Noise)
Prediction, evaluations, and mitigation of transit noise in the urban setting, including light rail, commuter rail, subway, streetcar, and bus transit

Underwater Acoustics (UW)

Automomous underwater vehicle acoustics (Joint with Engineering Acoustics)
Acoustic sensing navigation and communication concepts for autonomous underwater vehicles

Other Technical Events

Distinguished Lectures

The Technical Committees on Signal Processing in Acoustics, Architectural Acoustics, Psychological and Physiological Acoustics, and Animal Bioacoustics will sponsor a Distinguished Lecture titled ''Communication Acoustics'' presented by Jens Blauert of Ruhr University.

The Technical Committees on Noise and Architectural Acoustics will sponsor a Distinguished Lecture titled ''Noise: My 62 Years of It'' presented by Laymon N. Miller.

Student Design Competition in Architectural Acoustics

The Acoustical Society of America's Technical Committee on Architectural Acoustics and the National Council of Acoustical Consultants will sponsor a student design competition to be judged and displayed at the meeting in New York. The Student Design Competition is intended to encourage students in the disciplines of Architecture, Engineering, Physics, and other curriculums that involve building design and/or acoustics to express their knowledge of architectural acoustics and noise control in a schematic design of a facility in which acoustical considerations are of primary importance. The submitted designs will be judged by a panel of professional architects and acoustical consultants. An award of $1,000 will be made to the entry chosen as ''First Honors.'' Four awards of $500 each will be made to entries judged ''Commendation.''

Entries may be submitted by individual students or by teams of a maximum of three students. Entries must be received no later than 10 May 2004.

Students intending to enter must make their intentions known by sending an email on or before 1 April 2004.

Additional information may be found at the ASA website (http://asa.aip.org).

Meeting Program

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

Paper Copying Service

Note that the paper copying service may be replaced by a new online site. Details will be provided in the acceptance notices sent to authors and in updated meeting information posted on the ASA Home Page ⟨http://asa.aip.org⟩.

Tutorial Lecture

A tutorial presentation on Listening to the Acoustics in Concert Halls will be presented by Leo L. Beranek and David Griesinger on Monday, 24 May at 7:00 p.m.

Abstract

Listening to the acoustics in concert halls. Leo L. Beranek (975 Memorial Drive, Suite 804, Cambridge, MA 02138-5755) and David Griesinger (Lexicon, 3 Oak Park Drive, Bedford, MA 01730-1441)

How does acoustics affect the symphonic music performed in a concert hall? The lecture begins with an illustrated discussion of the architectural features that influence the acoustics. Boston Symphony Hall, which was built in 1900 when only one facet of architectural design was known, now rates as one of the world's great halls. How this occurred will be presented. Music is composed with some acoustical environment in mind and this varies with time from the Baroque to the Romantic to the Modern musical period. Conductors vary their interpretation according to the hall they are in. Well-traveled listeners and music critics have favorite halls. The lecture then presents a list of 58 halls rank ordered according to their acoustical quality based on interviews of music critics and conductors. Modern acoustical measurements made in these halls are compared with their rankings. Music recordings will be presented that demonstrate how halls that have different measured acoustical parameters sound. Photographs of a number of recently built halls are shown as examples of how these known acoustical factors have been incorporated into architectural design.

Lecture notes will be available at the meeting in limited supply. Those who register by 26 April are guaranteed receipt of a set of notes. The registration fee for this lecture is $5.00 for both preregistration and registration on-site. The lecture has been underwritten by an ASA Fellow. Please use the registration form in the call for papers or on the ASA Home Page at http://asa.aip.org to register for the tutorial session.

Special Meeting Features

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 proposal (E-mail preferred) to be received by 12 April to: Jolene Ehl, ASA, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502, Tel: 516-576-2359; Fax: 516-576-2377; E-mail: jehl@aip.org. The proposal should include your status as a student; whether you have submitted an abstract; whether you are a member of ASA; method of travel; if traveling by auto; whether you will travel alone or with other students; names of those traveling with you; and approximate cost of transportation.

Young Investigator Travel Grant

The Committee on Women in Acoustics is sponsoring a Young Investigator Travel Grant to help with travel costs associated with presenting a paper at the New York meeting. This award is designed for young professionals who have completed the doctorate in the past 5 years (i.e., not currently enrolled as a student), who plan to present a paper at the New York meeting. Each award will be of the order of $300. It is anticipated that the Committee will grant a maximum of three awards. Applicants should submit a request for support, a copy of the abstract they have submitted for the meeting, and a current resume/vita which provides information on their involvement in the field of acoustics and the ASA to: Melissa Epstein, Biomedical Sciences, Room 5A12, University of Maryland Dental School, 666 West Baltimore Street, Baltimore, MD 21201; Fax: 410-706-0193; E-mail: mae001@dental.umaryland.edu. Deadline for receipt of applications is 15 April.

Students Meet Members for Lunch

The Education Committee has established a program for students to meet one on one with members of the Acoustical Society over lunch. The purpose is to make it easier for students to meet and interact with members at meetings. Each lunch pairing is arranged separately. Students who wish to participate may sign up by contacting David Blackstock, University of Texas at Austin, Mechanical Engineering Dept., 1 University Station C2200, Austin TX 78712-0292; Tel.: 512-471-3145; Fax: 512-471-1045; E-mail: dtb@mail.utexas.edu. The information needed in order to arrange each pairing is the student's name, university, department, status (graduate student or undergraduate), research field, interests in acoustics, and days that are free for lunch. The sign-up deadline is one week before the start of the meeting, but an earlier sign-up is strongly encouraged. The cost of the meal is the responsibility of each participant.

75th Anniversary Celebration

75th Anniversary Celebratory Banquet

The Celebratory Banquet will be held on Tuesday, 25 May, at 6:30 p.m. A filmed presentation of the Society's ''Glorious Past'' featuring the past presidents will be presented. Some special awards will be given, and all present who have been members for 30 years or more will be honored. A musical presentation will be part of the entertainment. The banquet cost will be $65 per person for preregistration by 12 April and $75 thereafter including on-site registration at the meeting (cash bar). It is urged that tickets be reserved using the preregistration form in the call for papers or on the ASA Home Page at ⟨http://asa.aip.org⟩.

26 May—75th Anniversary Celebration Day

There will be no technical sessions on Wednesday, 26 May

The celebration will begin at 8:30 a.m. at City Center (55th Street between Ave. of the Americas and 7th Ave.). There is no admission charge, but **each person, including spouses and guests, must wear a meeting registration badge**. The festivities will start with the Plenary Session, at which major Society awards will be presented, special announcements made, the outgoing president honored, and the incoming president introduced. After a brief ''standing'' intermission, an extraordinary ''Celebratory Look Into The Future'' will be presented.

At about 11:30 a.m., a one-hour demonstration of an advanced virtual orchestra will be featured. Twenty-four loudspeakers with tonal quality to match the acoustical instrument being emulated will be located in the orchestra pit of the theater. A computer operator will have complete real-time control over orchestral dynamics and tempo. The computer operator and two singers will follow a professional conductor.

Immediately following the celebratory session, the College of Fellows will hold a reception for all past and new ASA Fellows in the Mezzanine Lobby of City Center (second floor). Photos will be taken for the ASA archives. Champagne and cheese (no lunch) will be served. The cost is $25 per attendee (plus one free guest per Fellow). The reception is expected to last until 1:30 p.m. Space is limited to 300 guests and Fellow counts are needed so please make your reservation using the preregistration form on the ASA Home Page ⟨http://asa.aip.org⟩ or in the printed Call for Papers

Premier Tour: ASA has chartered a Circle-Line cruise boat for a tour around Manhattan Island, with a tour guide, that lasts 3 hours and includes skyline views of: Battery Park, Ellis Island, Statue of Liberty, Wall Street, Brooklyn Bridge, Empire State Building, Chrysler Building, United Nations Buildings, Gracie Mansion, Yankee Stadium, New Jersey Palisades, George Washington Bridge, Grant's Tomb, and more. Tickets are $40 for adults and $25 for children, including bus transportation between hotel and pier; lead bus loads at 1:45 p.m. Box lunches are available for $17.

A tour bus to the Steinway Piano Factory will depart from City Center at 1:30 p.m. The Steinway Factory was recently featured in *The New York Times* and is a very special and historical site. Box lunches will be available on the bus for $17. The bus will return to the hotel at about 4:30 p.m. The event is sponsored by the Acoustical Society Foundation and the cost is $60. This tour is limited to 44 participants.

A tour to the Rose Planetarium at the Museum of Natural History will depart from the hotel at 3:30 p.m. This tour features attendance at a regularly scheduled show in the Museum followed by a one-hour ''behind the scenes'' private tour conducted by the chief audio engineer. He will demonstrate the most complex multichannel spatialized audio system ever built and equally complex video computer system. The bus will return to the hotel about 6:00 p.m. The event is sponsored by the Acoustical Society Foundation and the cost is $60. This tour is limited to 44 participants.

A walking tour to two of New York's finest organs with performances will be held on Wednesday afternoon. The tour will leave from the Sheraton at 2:00 p.m. and return at about 5:30 p.m. The cost of the tour is $60 per person.

A pipe organ concert will occur at the St. Thomas Episcopal Church, 53rd Street and 5th Avenue (a short 5–10 minute walk from hotel), 26 May, at 7:00 p.m. The concert is expected to last about one and one-half hours. There is no admission charge.

New York City tours leave from a stop across the street from the Sheraton every 20 minutes. The most interesting tours offer passes good for 24 hours, that allow one to get on and off the bus and stop to see specific sights or to eat. If one stays on the bus, the tours take 2 to 3 hours to complete. For details, refer to www.just-tours.net; or www.shop.viator.com, or www.allnewyorktours.com. (Cost by E-mail or at departure terminal $39 to $49 for upper or lower Manhattan tours.)

New York's famous Fifth Avenue shopping street is two blocks east from the hotel and features such stores as Bergdorf Goodman, Saks Fifth Avenue, Tiffany's, Lord and Taylor, and a multitude of fine shops from all countries.

BRING RAINCOAT AND UMBRELLA! Hoping for sunshine all week.

Social Event

On Thursday evening a social hour will be held to which all registrants are invited as guests of the Society. The event will take place at the Sheraton New York Hotel and Towers. Dress casually. Cash bar.

Transportation and Hotel Accommodations

Air Transportation

New York City is served by three airports: John F. Kennedy International (JFK), LaGuardia Airport (LGA), and Newark Liberty International Airport (EWR). Information for all three airports can be found on the Port Authority website, ⟨http://www.panynj.gov/aviation.html⟩ Please note that a three-day holiday weekend (29–31 May) follows the week of the meeting, and air and automobile traffic leaving New York on Friday, 28 May, is expected to be higher than usual.

Ground Transportation

Major car rental companies. Rental car counters are located adjacent to the airports' baggage claim areas. Note that car rental is not available at the terminals in Newark Airport. Use the AirTrain to parking lot P2 or P3. Parking at the Sheraton is $37 per day.

SuperShuttle—$17 from Newark and JFK; $15 from LGA, 1-800-258-3826.

New York Service Express Bus—Follow the signs for ground transportation for the New York Service Express Bus out of JFK and LaGuardia Airports. The bus operates from 6:15 a.m. to 11:10 p.m., every 15–30 minutes. The fare to Grand Central train station (42nd Street between Lexington and Fifth Avenues) is $10–$13. Transfers to the Sheraton New York Hotel and Towers (7th Avenue and 52nd Street) are available for an additional charge. Reservations for the return trip can be made by calling 212-875-8200.

Olympia Airport Express—Follow the signs for ground transportation for the Olympia Airport Express out of Newark Airport. The bus operates from 4:00 a.m. to 11:00 p.m., every 30–60 minutes. The fare to Grand Central train station (42nd Street between Lexington and Fifth Avenues) is $11. Transfers to the Sheraton New York Hotel and Towers (7th Avenue and 52nd Street) are available for an additional charge. Reservations for the return trip can be made by calling 212-964-6233.

Taxi (with normal traffic)—JFK; $35 plus tolls; LaGuardia; $19–22 plus tolls; Newark; $45–55.

Driving in Manhattan is not recommended.

Hotel Accommodations and Reservations

The meeting will be held at the Sheraton New York Hotel and Towers. Please make your reservations directly with the hotel and ask for one of the rooms being held for the Acoustical Society of America (ASA). The reservation cut-off date for the special discounted ASA rates is 1 May 2004; after this date, the conference rate will no longer be available. You must mention the Acoustical Society of America when making your reservations to obtain the special ASA meeting rates.

The Sheraton New York Hotel and Towers is located in midtown Manhattan a few blocks north of Times Square and the theater district. It is located on the east side of 7th Avenue between 52nd and 53rd Streets. The hotel features over 1,700 guest rooms, a complete health club with 24-hour access, 50-foot indoor pool, and seasonal sun deck. Guest services include in-room refreshments, in-room movies, cable TV, and coffee makers.

Sheraton New York Hotel and Towers
811 7th Avenue at 52nd Street
New York, NY 10019
Tel: 212-581-1000
Fax: 212-262-4410

Rates
Single: $199 plus tax
Double: $219 plus tax
A limited number of rooms are available at the government rate.

Room Sharing

ASA will compile a list of people who wish to share a hotel room and its cost. To be listed, send your name, telephone number, E-mail address, gender, smoker or nonsmoker preference, by 23 April to the Acoustical Society of America, preferably by E-mail: asa@aip.org or by postal mail to Attn.: Room Sharing, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502. The responsibility for completing any arrangements for room sharing rests solely with the participating individuals.

Weather

New York generally enjoys nice spring weather in May, with average high temperatures of 75 °F and average lows of 55 °F with the possibility of spring showers.

General Information

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: Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502, asa@aip.org.

Accompanying Persons Program

Spouses and other visitors are welcome at the New York meeting. The registration fee for accompanying persons is $75. A hospitality room for accompanying persons will be open at the Sheraton Hotel from 8:00 a.m. to 11:00 a.m. each morning throughout the meeting where information about activities in and around New York City will be provided.

New York City has five boroughs—the Bronx, Brooklyn, Manhattan, Queens, and Staten Island—that are linked by a series of bridges, tunnels, subways, and ferries. The ASA meeting hotel is located in Manhattan, which is an island; the Bronx is north of Manhattan and is on the mainland; Queens and Brooklyn are on the western tip of Long Island, which stretches east into the Atlantic Ocean. Staten Island is located in New York Bay to the southwest of Manhattan.

Getting around New York City is facilitated by 3,700 buses, 714 miles of subways, 12,000 taxis and limousines, and countless feeder roads. There are also ferries, helicopters, bicycles, and frequent Amtrak and commuter rail service. And don't forget your feet! NYC is a walking city—flat and much of it on a grid.

New York offers a wealth of entertainment options from aquariums, zoos, and museums, shopping and dining, and a broad range of entertainment including city tours, sports, arts and cultural events, and a multitude of activities especially for children. One of the best resources for tourist information, the Official Visitor Information Center, is a short walk from The Sheraton.

New York City's Official Visitor Information Center, 810 Seventh Ave. between 52nd & 53rd Streets, Tel: 212/484-1222; Hours: Monday to Friday

8:30 a.m. to 6 p.m.; Saturday, Sunday, and Holidays 9 a.m. to 5 p.m. New York City's Official Visitor Information Center is the city's official source for information on everything there is to do and see in New York City including culture, dining, shopping, sightseeing, events, attractions, tours, and transportation. The Center features free brochures, discount coupons to attractions and theaters, multilingual visitor information counselors, ATM, and MetroCard vending machine. Stop in, pick up brochures and discount coupons, find out what's going on, get directions...it's a great place to start your visit. Half-price tickets for theaters may be obtained on the day of the performance at the TKTS booth located at 7th Avenue and 47th Street.

Additional information about things to do in New York City during your visit is available online at http://www.nycvisit.com.

Information about New York theaters and concerts may be obtained from AmericanaTicketsNY, and tickets ordered in advance ⟨http://www.americanatickets.com⟩. The telephone number of their branch in the Sheraton Hotel is 212-397-0372.

Registration Information

The registration desk at the meeting will open on Monday, 24 May, at the Sheraton New York Hotel and Towers. To register use the forms in the printed Call for Papers or register online at ⟨asa.aip.org⟩. **If your registration is not received at the ASA headquarters by 26 April you must register on-site**.

Registration fees are as follows:

Category	Preregistration by 12 April	Registration after 12 April
Acoustical Society Members	$350	$400
Acoustical Society Members One-Day	$175	$200
Nonmembers	$400	$450
Nonmembers One-Day	$200	$225
Nonmember Invited Speakers	$75	$75
Students (with current ID cards)	$35	$35
Emeritus members of ASA (Emeritus status preapproved by ASA)	$75	$75
Accompanying Persons (Spouses and other registrants who will not participate in the technical sessions)	$75	$75

Nonmembers who simultaneously apply for Associate Membership in the Acoustical Society of America will be given a $50 discount off their dues payment for the first year (2004) of membership. (Full price for dues: $100.)

NOTE: A $25 PROCESSING FEE WILL BE CHARGED TO THOSE WHO WISH TO CANCEL THEIR REGISTRATION AFTER 12 APRIL.

75th Anniversary Committee

Leo L. Beranek, Co-Chair
Patricia K. Kuhl, Co-Chair
Anthony A. Atchley
Fredericka Bell-Berti
Richard H. Campbell
William J. Cavanaugh
Dan Clayton
Lawrence A. Crum
Christopher J. Jaffe
Francis Kirschner

147th Local Organizing Committee

Damian J. Doria, Co-Chair
Russell Johnson, Co-Chair
Gary W. Elko, Technical Program Chair
Katharine Sawicki, Assistant to the Chairs & A/V
Pam Brooks, Accompanying Persons Program
Konrad Kaczmarek, Signs
Dan Clayton, Technical Tours
Richard H. Campbell, Tours
Fredericka Bell-Berti, Education
Subha Maruvada, At Large
Jeffrey Ketterling, At Large

ASA Celebrates the 25th Anniversary of the F. V. Hunt Fellowship in Acoustics

In recognition of 25 years of successful operation of the F. V. Hunt Postdoctoral Fellowship Program, a commemorative luncheon was held during the recent meeting of the Society in Austin, Texas. More than half of the Fellows were present, along with Society officers, members of the Prizes and Special Fellowships Committee, and other individuals who have made valuable contributions to the effectiveness of the Program. It was a special treat to be joined by Dr. Thomas Hunt, son of the late ASA President, Gold Medalist, and Harvard University professor Fredrick V. Hunt, for whom the Program is named.

Included in a brochure prepared for the luncheon was a brief summary for each of the 25 Hunt Postdoctoral Research Fellows, along with a few words about Professor Hunt. These are being published in the Journal in order to provide a wider distribution of information about the outcomes of this Society program.

FIG. 1. Past Hunt Fellows and others involved in the Hunt Fellowship over the past 25 years at the celebration luncheon held in Austin.

1978–1979 · Steven L. Garrett

Dr. Garrett worked in the ultralow temperature group at the University of Sussex, UK. There he continued research on the nonlinear parametric generation of sound in superfluid helium, which was the subject of his doctoral thesis at the University of California, Los Angeles. Other activities included a survey of experimental studies in liquid helium up to that time, as well as the development of a Rayleigh disk experiment for use in superfluid helium 3. After that postdoctoral year he assumed a position in the Physics Department of the Naval Postgraduate School, Monterey, California, and he is currently United Technologies Professor of Acoustics in the Graduate Program in Acoustics at The Pennsylvania State University.

1979–1980 · Mary J. Osberger

After completion of graduate work at the City University of New York, Dr. Osberger remained there for her Hunt-supported postdoctoral year. Her research topic was the computer simulation of the speech of the deaf. Since that time she has worked at the Indiana University School of Medicine in Indianapolis, and she's now concerned with clinical studies toward improvement in cochlear implant technology, for Advanced Bionics in Sylmar, California.

1980–1981 · Cynthia A. Prosen

With a doctorate from the University of Michigan in Ann Arbor, Dr. Prosen benefited from her Hunt Fellowship support at Northwestern University in Evanston, Illinois. There, working under the supervision of Dr. Peter Dallos, she greatly expanded her research skills and established a productive, flexible auditory behavioral laboratory for future research projects. Her specific research study was of behavioral and electrophysiological measures of frequency encoding in the chinchilla. Her present position is as Associate Professor of Psychology at Northern Michigan University in Marquette.

1981–1982 · Daniel Rugar

After completion of graduate work at the University of California, Los Angeles, Dr. Rugar carried out postdoctoral research at Stanford University's Edward L. Ginzton Laboratory. He devoted most of his effort to the development of an ultrahigh resolution acoustic microscope. While he continues his affiliation with Stanford as a Consulting Professor of Applied Physics, his primary position now is Manager of Nanoscale Studies at the IBM Almaden Research Center, San Jose, California, where he is trying to extend magnetic resonance imaging to the atomic scale.

1982–1983 · Wesley N. Cobb

Dr. Cobb, following graduate study at The University of Texas at Austin and at Yale University, pursued his postdoctoral work with Professor Robert Apfel in Yale's Applied Mechanics Department. His research topic was development of a frequency-domain method for the prediction of the ultrasonic field patterns of pulsed, focused radiators. He subsequently held a number of industry positions, with focus on the development of ultrasonic instruments for oil-field and pipeline applications, and is currently Research Professor of Engineering, and Director of the Center for Ultrasound and NDE, at the University of Denver.

1983–1984 · Mark F. Hamilton

With a doctorate in Acoustics from Pennsylvania State University, following graduate research that was completed at The University of Texas at Austin, Dr. Hamilton enjoyed a postdoctoral year at the Mathematics Institute of the University of Bergen, Norway. There he worked with Professors Jacqueline and Sigve Tjøtta on theoretical research in nonlinear acoustics, particularly finite-amplitude effects in sound beams and waveguides. He is currently the Harry L. Kent, Jr. Professor in Mechanical Engineering at UT Austin, and Vice President-Elect of the ASA.

1984–1985 · Christine H. Shadle

Following graduate work on the acoustics of speech production at the Massachusetts Institute of Technology, Dr. Shadle spent her Hunt postdoctoral year at the Institute of Sound and Vibration Research at the University of Southampton, UK. There she extended her thesis research, working on the aeroacoustics of fricative consonants, and also worked on the aeroacoustics of laryngeal vibration, using flow visualization of a model of the vocal folds and tract. A second beneficial experience was provided by a NATO postdoctoral fellowship at the Department of Speech Communication and Music Acoustics of KTH, Stockholm, Sweden. She then accepted an academic position in the School of Electronics and Computer Science at the University of Southampton.

1985–1986 · Anthony A. Atchley

A Physics graduate from the University of Mississippi, Dr. Atchley carried out his postdoctoral research in Professor Robert Apfel's group at Yale University. His research there, concerned with cavitation nucleation, the detection of acoustic cavitation, and the dependence of the cavitation threshold on various parameters, resulted in several publications and presentations. After that Hunt Fellowship-supported year he joined the faculty of the Naval Postgraduate School in Monterey, California. He is presently Chair of the Graduate Program in Acoustics at The Pennsylvania State University, as well as Vice President of ASA.

1986–1987 · M. Christian Brown

Dr. Brown, who prepared at the University of Michigan, pursued postdoctoral research in auditory neuroanatomy and neurophysiology at the Massachusetts Eye and Ear Infirmary in Boston. His goal was to study the feedback system that conveys information from the brain to the inner ear, the olivocochlear efferent system. He currently holds a joint appointment as Associate Professor of Otology and Laryngology at the Massachusetts Eye and Ear Infirmary and at Harvard Medical School.

1987–1988 · Ian M. Lindevald

Upon completion of graduate studies in Physics at Case Western Reserve University, Dr. Lindevald spent his postdoctoral year with Dr. Ernst Terhardt at the Institute of Electroacoustics of the Technical University of Munich, Germany. His primary experimental research was an investigation of the perceived pitch of spectrally dense organ tones in the reverberantly decaying sound field of a church. A number of additional projects also were completed. He now is on the Science Division faculty of Truman State University in Kirksville, Missouri.

1988–1989 · Elizabeth C. Oesterle

Dr. Oesterle carried out Ph.D. thesis research in physiological acoustics with Dr. Peter Dallos of Northwestern University, and as a Hunt Fellow she continued research in the same area at the University of Washington, Seattle with Dr. Edwin Rubel. She investigated the hypothesis that proteins known to be present in CNS glia are also present in inner-ear supporting cells. She remained at the University of Washington following the fellowship year, with interests focused on anatomy and physiology of the inner ear and hair cell regeneration, and is now Research Associate Professor in the Department of Otolaryngology—Head and Neck Surgery.

1989–1990 · E. Carr Everbach

Following his graduate studies at Yale University, Dr. Everbach carried out postdoctoral research in the Rochester Center for Biomedical Ultrasound at the University of Rochester, New York. The main goal of his research was to determine the physical mechanisms of kidney stone and gallstone destruction by acoustic lithotripsy. Collaborative efforts with several scientists and MDs led to a number of successful research studies. Dr. Everbach has continued to actively participate in University of Rochester research while holding a teaching position in the Engineering Department at Swarthmore College in Pennsylvania.

1990–1991 · Kenneth A. Cunefare

Upon completion of his graduate studies at The Pennsylvania State University, Dr. Cunefare spent his Hunt postdoctoral year in residence at the Institute for Technical Acoustics, Technical University of Berlin, Germany, working with Herr Dr. Professor Manfred Heckl. His research topic was the investigation of the concept of exterior acoustic radiation modes, the idea that exterior acoustic fields could be decomposed in an orthogonal expansion (via an eigenvalue problem or a singular value decomposition), where the eigenvalues correspond to radiation efficiencies and the corresponding eigenvectors correspond to particular vibration patterns on the surface of the body of interest. He now is on the faculty of the School of Mechanical Engineering at the Georgia Institute of Technology, Atlanta.

1991–1992 · Gregory J. Sandell

With strong interest in musical acoustics and a degree from Northwestern University, Dr. Sandell traveled to the University of California, Berkeley for his Hunt Fellowship year. There he investigated auditory science and musical sound processing with David Wessel. Subsequently he held additional postdocs, for 2 years in the Experimental Psychology Department at

the University of Sussex in Brighton, England, and for 3 years at the Parmly Hearing Institute of Loyola University in Chicago, Illinois, before turning to employment in the private sector.

1992–1993 · Quan Qi

Dr. Qi carried out both graduate and postdoctoral research in the Department of Theoretical and Applied Mechanics at the University of Illinois at Urbana–Champaign. His main focus was investigation of the mechanisms of ultrasonic cleaning of small particles without using cavitation and a chemical detergent. He also studied the effect of compressibility on acoustic streaming near a plane boundary, and ultrasonic cleaning of compliant surfaces such as skin. Following experience in the Department of Mechanical Engineering, the University of Colorado at Boulder, he is currently at the Hewlett-Packard Company in Corvallis, Oregon.

1993–1994 · Charles E. Bradley

Following completion of graduate research in nonlinear acoustics at The University of Texas at Austin, Dr. Bradley traveled for his postdoctoral experience to the University of California at Berkeley. There he focused his attention on the possible use of acoustic streaming for cooling microelectromechanical systems, including miniature transducers. After a second postdoctoral year at UC, he accepted a research position in California with Siemens Medical Solutions Ultrasound Division.

1994–1995 · T. Douglas Mast

Armed with a Pennsylvania State University doctorate, Dr. Mast spent his Hunt Fellowship year at the University of Rochester Diagnostic Ultrasound Research Laboratory in Rochester, New York. There he collaborated with Professor Robert Waag and others on theoretical, computational, and experimental investigations of acoustic scattering from soft tissues. These investigations had the general goal of improved methods for ultrasonic diagnosis of disease. The specific motivation for the research was to develop improved quantitative and high-resolution diagnostic techniques for use with next-generation ultrasonic apparatus. Dr. Mast is currently working for Ethicon Endo-Surgery in Cincinnati, Ohio, a Johnson & Johnson company specializing in minimally invasive surgical devices.

1995–1996 · Robin O. Cleveland

After completion of his doctoral research at the University of Texas at Austin, Dr. Cleveland worked with Professor Lawrence Crum at the Applied Physics Laboratory (APL) of the University of Washington in Seattle. His task was to carry out research in extracorporeal shock wave lithotripsy—the medical procedure whereby shock waves are used to break up kidney stones. The specific original goal of the fellowship year was to model the propagation of lithotripsy shock waves through tissue. He remained at APL a second year as a full-time researcher, and is now a member of the faculty of the Department of Aerospace and Mechanical Engineering at Boston University.

1996–1997 · Mark A. Hasegawa-Johnson

Building upon his graduate studies at the Massachusetts Institute of Technology, Dr. Hasegawa-Johnson split his postdoctoral year between Cambridge and Los Angeles. He worked with Professor Abeer Alwan at UCLA and Professor Ken Stevens at MIT to study the relationships among articulatory dynamics, vowel formant dynamics, and vowel spectral dynamics. Their observations from the field of speech production led to exploration of a new model of speech recognition, in which the dependence of acoustic spectra on phoneme identity is mediated by the dynamics of two to four continuous-valued hidden state variables. Dr. Hasegawa-Johnson is currently at the Beckman Institute of The University of Illinois in Urbana.

1997–1998 · James J. Finneran

Following graduate study at Ohio State University, Dr. Finneran worked with Dr. Sam Ridgway at the U.S. Navy Marine Mammal Program, Space and Naval Warfare Systems Center (SSC), San Diego. Accomplishments during the postdoctoral year included development of a computer-controlled system for archiving the facility's rather large collection of magnetic tape-based data, and collaboration on a marine mammal temporary threshold shift project. Dr. Finneran has remained at SSC San Diego, first as an NRC Postdoctoral Research Fellow and now as a full-time employee studying the hearing capabilities of, and the effects of sound on, marine mammals.

1998–1999 · Lily M. Wang

Dr. Wang completed her graduate studies at Pennsylvania State University, and spent her postdoctoral year with Dr. Anders Gade in the Acoustic Technology Department at the Technical University of Denmark. Her research efforts were focused on investigating the phenomenon of spatial impression in concert halls, and they sought to clarify which objective measure (and therefore physical explanation) best matched the subjective perception of spatial impression. She currently is responsible for the Architectural Engineering Program at the University of Nebraska–Lincoln in Omaha.

1999–2000 · Penelope Menounou

With an F.V. Hunt Postdoctoral Research Fellowship, Dr. Menounou continued theoretical studies, in the Mechanical Engineering Department of the University of Texas at Austin, on the propagation of finite-amplitude noise. Her goal was to predict the effect of nonlinear propagation distortion of the power spectral density of a given noise source, when the time waveform itself is not known, only its power spectral density. A new method was developed that predicted the combined effect of nonlinearity and thermoviscous attenuation directly on the power spectral density of a given noise source. She has continued research on traffic noise propagation and diffraction in this department.

2000–2001 · James C. Lacefield

Following doctoral training at Duke University, and with the support of the F. V. Hunt Fellowship, Dr. Lacefield investigated methods for the estimation and compensation of medical ultrasound focusing errors (i.e., aberration) produced by propagation through heterogeneous tissue. This research was performed at the University of Rochester in collaboration with Professor Robert C. Waag. Dr. Lacefield is now a faculty member in the Departments of Electrical and Computer Engineering and Medical Biophysics, University of Western Ontario, London, Ontario, Canada.

2001–2002 · Chao-Yang Lee

With a solid background in the cognitive and linguistic sciences at Brown University, Dr. Lee carried out postdoctoral work in the Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology. The broad goal was to delineate how acoustic properties of speech are processed, based on knowledge of articulatory–acoustic–linguistic relationships, to access lexical representations. In particular, he explored whether the approach of lexical access based on detecting acoustic landmarks and extracting phonetic features, proposed by Professor Ken Stevens at MIT, could be extended to Mandarin Chinese, a tone language in which fundamental frequency variations are used for lexical distinctions. He now is on the faculty of the School of Hearing, Speech and Language Sciences at Ohio University, Athens, Ohio.

2002–2003 · Constantin-C. Coussios

Dr. Coussios completed his graduate studies at the University of Cambridge, England in biomedical engineering. His support as a Hunt Postdoctoral Research Fellow began late in 2002, and he worked with Professor Ronald Roy at Boston University through the Fall 2003 semester. His research topic involves the modeling of the cavitation bubble field in tissue-mimicking media. He has accepted a position as Lecturer in Biomedical Engineering and Fellow of Magdalen College at the University of Oxford, England.

2003–2004 · Tyrone M. Porter

The most recent F. V. Hunt Fellowship was awarded to Dr. Porter, who received his Ph.D. in bioengineering at the University of Washington in Seattle. He recently began his postdoctoral appointment at the University of

Cincinnati, Ohio, where he is working with Professor Christy Holland. His research interest is in the use of ultrasound to image blood clots and to assist in the targeted delivery of drugs.

Frederick Vinton Hunt · 1905–1972

F. V. Hunt, known from a young age as Ted, grew up in Barnesville, Ohio and earned AB and BEE degrees from Ohio State University. He enrolled in the graduate school of Harvard University in 1925 and remained there until his retirement, a year before his death. Following completion of doctoral dissertations in both engineering and physics, he made rapid progress through the academic ranks that culminated in two endowed chairs: Gordon McKay Professor of Applied Physics and Rumford Professor of Physics. He thrived on the opportunity to interact in Cambridge with a broad group of creative, distinguished scientists, with whom he sharpened his incisive mind and expanded his vision. Throughout his days he exhibited a remarkable gift for teaching, a brilliant flair for practical invention and development, and a skill for scientific administration.

Professor Hunt spoke of three periods in his professional life: before 1941, when he did things himself; the war years, in which he did things through a large research organization; and after World War II, when he stimulated his students to do things. His graduate study included research on room acoustics and development of electronic apparatus that made that research possible. Subsequently he developed lightweight phonograph pickups, wrote a textbook on electroacoustics, and contributed as advisor toward a continuing set of Navy concerns. He supervised graduate student research over a wide range of topics, from cavitation in liquids to signal processing. During World War II he organized and directed the Harvard Underwater Sound Laboratory, funded by the US government and with the objective to improve existing apparatus (as well as conceive and design new approaches) for the detection and localization of submarines. As was stated in the encomium for the ASA Pioneers of Underwater Acoustics Medal that he received in 1965, this diversion of his energy led to his making ''pioneering contributions to underwater acoustics as a scientist, innovator, teacher and administrator; and particularly [for his] unceasing efforts directed toward greater scientific understanding and more effective exploitation of sound in the sea.''

A member of the Acoustical Society of America as early as 1931, Ted Hunt served the Society as President and on its Councils and Committees, helping to shape the directions it has taken. He contributed his keen wit, humor, and pithy remarks to technical sessions in all but one of ASA's meetings during the remainder of his life.

The availability of the F. V. Hunt Postdoctoral Research Fellowship to students in all technical areas of the Society is consistent with Professor Hunt's deep concerns for the entire discipline. His award of the Society's Gold Medal in 1969 was accompanied by the citation: ''For his extensive contributions to the science and technology of acoustics in the fields of architecture, engineering, and signal processing; for his creative leadership in underwater sound and its application to the security of our nation; for his inspirational guidance as a teacher and his stimulating leadership as a laboratory administrator; and for his dedicated service to the Society as an author, lecturer, counselor, and officer.''

During the final decade of his teaching career he recognized that it was no longer possible for one person to remain up-to-date in so many technical areas, but he retained his interest in Acoustics as a whole. He never lost his enthusiasm, and he remained loyal and encouraging to all those who had been privileged to be his students, postdocs, colleagues, and friends. It is entirely fitting that this fellowship honors him and continues to bring his life to the attention of new generations of acousticians.

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.

2004

24–28 May	75th Anniversary Meeting (147th Meeting) of the Acoustical Society of America, New York, NY [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: http://asa.aip.org].
12–14 July	Noise-Con 2004, Baltimore, MD [Institute of Noise Control Engineering of the USA, Inc., INCE/USA Business Office, 212 Marston Hall, Iowa State Univ., Ames, IA 50011-2153; Tel.: 515-294-6142; Fax: 515-294-3528; E-mail: ibo@inceusa.org; WWW: http://www.inceusa.org/NoiseCon04call.pdf]
3–7 Aug.	8th International Conference of Music Perception and Cognition, Evanston, IL [School of Music, Northwestern Univ., Evanston, IL 60201; WWW: www.icmpc.org/conferences.html].
20–24 Sept.	ACTIVE 2004—The 2004 International Symposium on Active Control of Sound and Vibration, Williamsburg, VA (INCE Business Office, Iowa State Univ., 212 Marston Hall, IA 50011-2153 Fax: 515 294 3528; E-mail: ibo@ince.org; WWW: inceusa.org).
15–19 Nov.	148th Meeting of the Acoustical Society of America, San Diego, CA [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: http://asa.aip.org].

Cumulative Indexes to the Journal of the Acoustical Society of America

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

Some indexes are out of print as noted below.

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

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

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

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

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

Volumes 36–44, 1964–1968: Contemporary Literature. Classified by subject and indexed by author. Pp. 1060. Out of Print.

Volumes 45–54, 1969–1973: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 540. Price: $20 (paperbound); ASA members $25 (clothbound); Nonmembers $60 (clothbound).

Volumes 55–64, 1974–1978: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 816. Price: $20 (paperbound); ASA members $25 (clothbound); Nonmembers $60 (clothbound).

Volumes 65–74, 1979–1983: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 624. Price: ASA members $25 (paperbound); Nonmembers $75 (clothbound).

Volumes 75–84, 1984–1988: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 625. Price: ASA members $30 (paperbound); Nonmembers $80 (clothbound).

Volumes 85–94, 1989–1993: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 736. Price: ASA members $30 (paperbound); Nonmembers $80 (clothbound).

Volumes 95–104, 1994–1998: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 632. Price: ASA members $40 (paperbound); Nonmembers $90 (clothbound).

ACOUSTICAL NEWS—INTERNATIONAL

Walter G. Mayer
Physics Department, Georgetown University, Washington, DC 20057

International Meetings Calendar

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

March 2004

17–19 **Spring Meeting of the Acoustical Society of Japan**. Atsugi, Japan. (Fax: +81 3 5256 1022; Web: wwwsoc.nii.ac.jp/asj/index-e.html)

22–25 **Joint Congress of the French and German Acoustical Societies (SFA-DEGA)**, Strasbourg, France. (Fax: +33 1 48 88 90 60; Web: www.sfa.asso.fr/cfa-daga2004)

23–26 **International Conference: Speech Prosody 2004**, Nara, Japan. K. Hirose, School of Frontier Sciences, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; Fax: +81 3 5841 6648; Web: www.gavo.t.u-tokyo.ac.jp/sp2004)

31–3 **International Symposium on Musical Acoustics (ISMA2004)**, Nara, Japan. (Fax: +81 77 495 2647; Web: www2.crl.go.jp/jt/a132/isma2004)

April 2004

5–9 **18th International Congress on Acoustics (ICA2004)**, Kyoto, Japan. (Fax: +81 66 879 8025; Web: www.ica2004.or.jp)

11–13 **International Symposium on Room Acoustics (ICA2004 Satellite Meeting)**, Hyogo, Japan. (Fax: +81 78 803 6043; Web: rad04.iis.u-tokyo.ac.jp)

25–30 ***9th Meeting of the European Society of Sonochemistry**, Badajoz, Spain. (European Society of Sonochemistry, Avenida de Europa 3, 06004 Badajoz, Spain; Web: www.ess9.com)

27–28 ***Advanced Metrology for Ultrasound in Medicine**, Teddington, UK. (Fax: +44 20 8943 6735; Web: www.amum2004.npl.co.uk)

May 2004

8–10 **116th AES Convention**, Berlin, Germany. (Web: aes.org/events/116)

10–12 **Tenth AIAA/CEAS AeroAcoustics Conference**, Manchester, UK. (Web: www.aiaa.org)

17–21 **International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)**, Montréal, Canada. (Web: www.icassp2004.com)

June 2004

6–9 **13th International Conference on Noise Control**, Gdynia, Poland. (Web: www.ciop.pl/noise_04)

8–10 ***Transport Noise and Vibration 2004**, St. Petersburg, Russia. (Web: webcente.ru/~eeaa/tn04)

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

July 2004

5–8 **7th European Conference on Underwater Acoustics (ECUA 2004)**, Delft, The Netherlands. (Fax: +31 70 322 9901; Web: www.ecua2004.tno.nl)

5–8 **Eleventh International Congress on Sound and Vibration (ICSV11)**, St. Petersburg, Russia. (Web: www.iiav.org)

11–16 **12th International Symposium on Acoustic Remote Sensing (ISARS)**, Cambridge, UK. (Fax: +44 161 295 3815; Web: www.isars.org.uk)

August 2004

23–27 **2004 IEEE International Ultrasonics, Ferroelectrics, and Frequency Control 50th Anniversary Conference**, Montréal, Canada. (Fax: +1 978 927 4099; Web: www.ieee-uffc.org/index2-asp)

22–25 **Inter-noise 2004**, Prague, Czech Republic. (Web: www.internoise2004.cz)

30–1 **Low Frequency 2004**, Maastricht, The Netherlands. (G. Leventhall, 150 Craddlocks Avenue, Ashtead, Surrey KT 21 1NL, UK; Web: www.lowfrequency2004.org.uk)

September 2004

13–16 **Subjective and Objective Assessment of Sound**, Poznań, Poland. (Institute of Acoustics, Adam Mankiewicz University, Poznań, Poland. Web: www.ia.amu.edu.pl/index.html)

13–17 **4th Iberoamerican Congress on Acoustics, 4th Iberian Congress on Acoustics, 35th Spanish Congress on Acoustics**, Guimarães, Portugal. (Fax: +351 21 844 3028; Web: www.spacustica.pt/novidades.htm)

14–16 ***International Conference on Sonar Signal Processing and Symposium on Bio-Sonar Systems and Bioacoustics**, Loughboro, UK. (Fax: +44 1509 22 7053 [c/o D. Gordon]; Web: ioa2004.lboro.ac.uk)

15–17 **26th European Conference on Acoustic Emission Testing**, Berlin, Germany. (DGZIP, Max-Planck-Str. 26, 12489 Berlin, Germany; Web: www.ewgae2004.de)

20–22 **International Conference on Noise and Vibration Engineering (ISMA2004)**, Leuven, Belgium. (Fax: +32 16 32 29 87; Web: www.isma-isaac.be/fut_conf/default_en.phtml)

20–22 ***9th International Workshop "Speech and Computer" (SPECOM'2004)**, St. Petersburg, Russia. (Web: www.spiiras.nw.ru/speech)

28–30 **Autumn Meeting of the Acoustical Society of Japan**, Naha, Japan. (Fax: +81 3 5256 1022; Web: wwwsoc.nii.ac.jp/asj/index-e.html)

October 2004

4–8 **8th Conference on Spoken Language Processing (INTERSPEECH)**, Jeju Island, Korea. (Web: www.icslp2004.org)

6–8 ***Acoustics Week in Canada**, Ottawa, ON, Canada. (J. Bradley, NRC Institute for Research on Construction [Acoustics Section], Ottawa, Ontario, K1A 0R6; Fax: +1 613 954 1495; Web: caa-aca.ca)

November 2004

4–5 **Autumn Meeting of the Swiss Acoustical Society**, Rapperswil, Switzerland. (Fax: +41 419 62 13; Web: www.sga-ssa.ch)

May 2005

16–20 **149th Meeting of the Acoustical Society of America**, Vancouver, B. C. Canada. (ASA, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502 USA; Fax: +1 516 576 2377; Web: asa.aip.org)

August 2005

6–10 ***Inter-Noise**, Rio de Janeiro, Brazil. (Web: www.internoise2005.ufsc.br)

28–2 **EAA Forum Acusticum Budapest 2005**, Budapest, Hungary. (I. Bába, OPAKFI, Fö u. 68, Budapest 1027, Hungary; Fax: +36 1 202 0452; Web: www.fa2005.org)

September 2005

4–8 **9th Eurospeech Conference (EUROSPEECH'2005)**, Lisbon, Portugal. (Fax: +351 213145843; Web: www.interspeech2005.org)

June 2006

26–28 **9th Western Pacific Acoustics Conference (WESPAC 9)**, Seoul, Korea. (Web: www.wespac8.com/WespacIX.html)

September 2007

2–7 **19th International Congress on Acoustics (ICA2007)**. Madrid, Spain. (SEA, Serrano 144, 28006 Madrid, Spain; Web: www.ia.csic.es/sea/index.html)

Preliminary Announcements

September 2005

5–9 ***Boundary Influences in High Frequency, Shallow Water Acoustics**, Bath, UK. (Details to be announced later)

June 2008

23–27 **Joint Meeting of European Acoustical Association (EAA), Acoustical Society of America (ASA), and Acoustical Society of France (SFA)**. Paris, France. (Details to be announced later)

BOOK REVIEWS

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

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

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

An Introduction to Underwater Acoustics

Xavier Lurton

Springer, New York, 2002
347 pp. Price: $89.95 (hardcover), ISBN 3540429670.

This book is a general survey of Underwater Acoustics, intended to make the subject ''as easily accessible as possible, with a clear emphasis on applications.'' In this the author has succeeded, with a wide variety of subjects presented with minimal derivation, that task being deliberately left to other sources. A book of this breadth must be compared to Urick's classic, *Principles of Underwater Sound*. A joke among some theoretically inclined underwater acousticians is that, while none will admit to using this ''cookbook,'' all have copies within reach just in case the real world of noise and reverberation intrudes. Lurton's book provides coverage of several subjects not found in Urick, with material on nonmilitary uses of underwater acoustics and deeper treatment of signal processing, including beamforming. There is an emphasis on technology and on intuitive physical explanation, with somewhat fewer of the curves, tables, and nomograms marked by dog-eared pages in Urick. Still, there are many useful formulas and curves, such as the Francois–Garrison absorption equations, the Chen–Millero sound-speed formulas, the Knudson noise curves, expressions for the strengths of simple targets, etc.

This book seems best suited to a graduate survey course for nonspecialists where it might, in the author's words, ''arouse vocations for acoustics, both in students and in scientists...'' For new practitioners, it would serve as an introduction and portal to the primary literature. For example, the treatment of sound propagation in the ocean waveguide is quite brief and would require use of cited references in many applications. In compensation, simple illustrative models, such as one employing multiple image sources, are presented with emphasis on ray methods. Outlines are given of the parabolic equation and modal methods, and these methods are supported by short appendixes, as are several other of the more technical subjects. The treatment of scattering by the ocean surface and bottom is balanced, brief, yet self-contained, and the chapter on acoustic mapping systems with accompanying color plates offers a wide-ranging description of several important techniques. A section on the use of decibels will be welcomed by students wondering whether a factor of 10 or a factor of 20 should precede the logarithm, although the somewhat cavalier use of units preferred in underwater acoustics is followed throughout. Thus, source voltage sensitivity is expressed in dB *re*: 1 μPa/1 m/1 V, implying that one should multiply by range to get pressure, when division is required. Similarly, interface scattering strength is the decibel equivalent of a dimensionless quantity, but is expressed as dB *re*: 1 m^2. Perhaps it's not the job of a book at this level to rectify these erroneous conventions, but they do little to advance learning. A short section on real, analytic, and quadrature signal representations is modern and useful, although more consistent notation would help the student sort out the differences. The book has numerous typographical errors and difficulties in translation. Some translation problems, such as ''writes as'' for ''can be written as'' are innocuous, but some, such as ''swell spectrum'' for ''wave spectrum,'' could cause difficulty. If these problems are fixed in a future edition, Lurton's handsome book will be a full realization of the author's goal of making his subject accessible to a wide audience.

DARRELL R. JACKSON
Applied Physics Laboratory
University of Washington
Seattle, Washington 98105

 0001-4966/2004/115(2)/443/1/$20.00

REVIEWS OF ACOUSTICAL PATENTS

Lloyd Rice
11222 Flatiron Drive, Lafayette, Colorado 80026

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

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
MARK KAHRS, *Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261*
HASSAN NAMARVAR, *Department of BioMed Engineering, University of Southern California, Los Angeles, California 90089*
DAVID PREVES, *Starkey Laboratories, 6600 Washington Ave. S., Eden Prarie, Minnesota 55344*
DANIEL R. RAICHEL, *2727 Moore Lane, Fort Collins, Colorado 80526*
CARL J. ROSENBERG, *Acentech Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
WILLIAM THOMPSON, JR., *Pennsylvania State University, University Park, Pennsylvania 16802*
ERIC E. UNGAR, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
ROBERT C. WAAG, *Univ. of Rochester, Department of Electrical and Computer Engineering, Rochester, New York 14627*

6,606,312

43.28.Bj ANALOG RADIO SYSTEM WITH ACOUSTIC TRANSMISSION PROPERTIES

Grant McGibney, assignor to Telecommunications Research Laboratories
12 August 2003 (Class 370/345); filed 25 March 1999

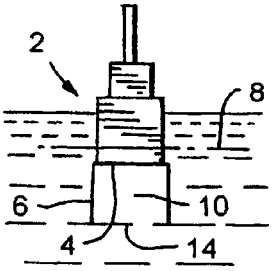
Multipath distortion in FM radio reception is audible and objectionable, yet for two people conversing in a typical room the acoustic multipath distortion is considered normal. ''The purpose of this invention is to coerce the radio medium to behave like the acoustic medium when carrying voice.'' Speech segments are digitized, then broadcast at a much higher sampling rate. This compresses the signal in time and expands it in frequency. The process is reversed in the radio receiver. ''After decompression, the signal from the radio channel sounds like a natural acoustic signal so there is no need for complex digital signal processing.''—GLA

6,606,279

43.30.Gv SUBMERGENCE DETECTION FOR ELECTRO ACOUSTIC TRANSDUCERS

Ashley Preston, assignor to Milltronics Limited
12 August 2003 (Class 367/99); filed in the United Kingdom 25 February 1998

A method for detecting the actual submergence of an electroacoustic transducer into a body of fluid is discussed. The transducer **2** has its radiating face **4** surrounded by a skirt **6** that encloses a volume of air **10**, even under submergence conditions. High-frequency pulses emitted by the transducer reflect, of course, from the air–water interface and the echoes are sensed by some receiver (not shown). The novelty of this patent seems to reside in how the echo signals are processed to determine incipient submersion of the transducer.—WT

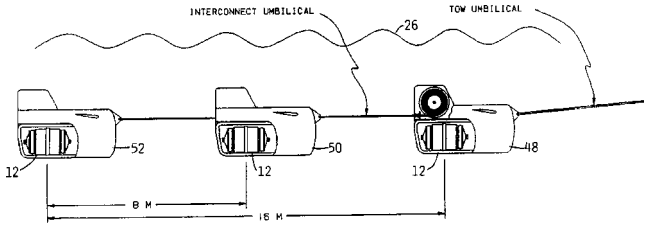


6,606,958

43.30.Yj TOWED ACOUSTIC SOURCE ARRAY SYSTEM FOR MARINE APPLICATIONS

John V. Bouyoucos, assignor to Hydroacoustics Incorporated
19 August 2003 (Class 114/242); filed 14 June 2000

A towed array comprises a set of individual positively buoyant tow bodies (**48**, **50**, **52** shown in the figure) each of which contains an acoustic source **12** such as an air gun or a hydraulic vibrator unit. Each tow body also contains a reel that stores an interconnect umbilical cable within its vertical stabilizer surface (shown in cutaway in the right-hand tow body). The reel is controllable from the tow ship through the tow umbilical cable. Hence the positioning, i.e., separation, of the tow bodies can be varied both for operational purposes as well as to facilitate their retrieval and storage. Other control surfaces on each tow body adjust its depth and transverse spatial positions.—WT



6,607,050

43.30.Yj INTEGRATED OCEAN BOTTOM TOWED ARRAY FOR FOUR-COMPONENT SEISMIC DATA ACQUISITION

Guoxin He *et al.*, assignors to China National Petroleum Corporation
19 August 2003 (Class 181/110); filed in China 26 April 2000

An ocean bottom towed seismic array consists of a hydrophone section (a piezoceramic accelerator sensor), a geophone section containing three mutually perpendicular velocity sensors all on the same gimbal unit, an electronics section, and cable joints, all connected in longitudinal arrangement on an underwater cable.—WT

6,595,002

43.35.Ud VIBRATION REDUCTION IN A COMBUSTION CHAMBER

Wolfgang Weisenstein, assignor to ABB Schweiz AG
22 July 2003 (Class 60/772); filed in the European Patent Office 1 May 2001

In annular combustion chambers of turbomachines in which burners are arranged around the circumference there occur self-excited pressure oscillations due to the interaction of the thermoacoustic pressure pulsations and local heat release from the burners. The resulting circumferential pressure oscillations may become very severe. In order to limit these oscillations, a number of modulatable burners are designed so that their fuel mass flow is modulated at frequencies different from the natural frequencies and from the instability frequencies (determined by the geometry and thermophysical conditions) of the combustion chamber.—EEU

6,545,947

43.35.Yb ACOUSTIC MATCHING MATERIAL, METHOD OF MANUFACTURE THEREOF, AND ULTRASONIC TRANSMITTER USING ACOUSTIC MATCHING MATERIAL

Hideki Morozumi *et al.*, assignors to Matsushita Electric Industrial Company, Limited
8 April 2003 (Class 367/152); filed in Japan 12 November 1999

Materials to match the acoustic impedance of solid transducer surfaces to that of air are obtained by using a large number of fine pieces, such as small hollow glass spheres, that are bonded to each other at contact points or encapsulated in a matrix.—EEU

6,607,491

43.35.Yb ULTRASONIC PROBE

Shohei Sato, assignor to Aloka Company, Limited
19 August 2003 (Class 600/459); filed in Japan 27 September 2001

The principal feature of this ultrasonic probe is its incorporation of a sonic speed control element. The probe consists of a transducer element, a sonic speed control element, and an acoustic-impedance matching layer. The sonic speed control element functions to control the sonic speed of the ultrasonic waves traveling therein, and it has an inclined characteristic in the speed control effect whereby that effect changes continuously along the direction of the travel of the ultrasonic waves. The specific acoustic impedance of the speed control element at the end nearest living tissue corresponds to the specific impedance of the acoustic matching layer at the end away from the tissue. In the preferred embodiment the acoustic matching layer is formed by layering two or three or more overlapped members. The boundary between two adjacent members has a shape featuring a series of hills and valleys. The crossing angle between the propagation direction of the ultrasonic waves and the boundary is set so that a predetermined condition is satisfied.—DRR

6,545,945

43.35.Zc MATERIAL CLASSIFICATION APPARATUS AND METHOD

David D. Caulfield, assignor to Ocean Data Equipment Corporation
8 April 2003 (Class 367/87); filed 26 February 2001

Materials in a container (such as a suitcase) or buried in the ground (e.g., explosives or pipelines) are identified by the use of electromagnetic or acoustic sources and receivers. Energy pulses of different frequencies are directed at the area of interest and the received signals are analyzed to determine the rate of change of energy absorption with frequency. This information can then be compared with a database of known objects.—EEU

6,552,803

43.35.Zc DETECTION OF FILM THICKNESS THROUGH INDUCED ACOUSTIC PULSE-ECHOS

Haiming Wang *et al.*, assignors to KLA–Tencor Corporation
22 April 2003 (Class 356/503); filed 17 August 1999

Measurement of the thickness of a film in microelectronics manufacturing is accomplished by directing laser pulses at the film surface. The resulting acoustic pulse is reflected from the interface between the film and the substrate, and is again reflected from the free surface. Heterodyne interferometry is used to measure the time interval between the first and second reflections, from which the film thickness is determined.—EEU

6,554,826

43.35.Zc ELECTRO-DYNAMIC PHASED ARRAY LENS FOR CONTROLLING ACOUSTIC WAVE PROPAGATION

Dana L. Deardorff, assignor to TxSonics–Limited
29 April 2003 (Class 606/27); filed 21 April 2000

This ultrasonic lens consists of a planar array of cells, each element of which may be controlled separately so that incident ultrasonic energy can be focused at desired points without the need for changing lenses.—EEU

6,566,886

43.35.Zc METHOD OF DETECTING CRYSTALLINE DEFECTS USING SOUND WAVES

Terri A. Couteau *et al.*, assignors to Advanced Micro Devices, Incorporated
20 May 2003 (Class 324/514); filed 28 March 2001

In this method for inspecting integrated circuit structures, the structure's natural frequency is determined. Then the structure is immersed in a liquid and a series of sound pulses are introduced in order to produce collapsing bubbles near the structure. These bubbles produce pulses in a frequency range that includes the structure's natural frequency, causing the structure to resonate. Thereafter the structure is examined for damage, permitting identification of flaws early in the manufacturing process.—EEU

6,596,239

43.35.Zc ACOUSTICALLY MEDIATED FLUID TRANSFER METHODS AND USES THEREOF

Roger O. Williams *et al.*, assignors to EDC Biosystems, Incorporated
22 July 2003 (Class 422/100); filed 12 December 2000

A source generates an acoustic wave that is focused by a lens prior to propagating through a layer of coupling fluid. The wave is then transmitted through a containment structure so as to come to a focus near the surface of a thin pool of source fluid, ejecting a droplet of that fluid. The containment structure can be supported on a stage that can be moved via computer-controlled actuators.—EEU

6,598,958

43.35.Zc LIQUID EJECTOR

Jyunichi Aizawa *et al.*, assignors to Mitsubishi Denki Kabushiki Kaisha
29 July 2003 (Class 347/46); filed in Japan 30 November 2000

Ejection of a liquid, as in an ink jet print head, is produced by impingement of an acoustic wave on a thin film of the liquid. A transducer injects an essentially longitudinal wave in a waveguide that has an appropriate shape, such as a paraboloid of revolution, which focuses the energy on the liquid film.—EEU

6,603,118

43.35.Zc ACOUSTIC SAMPLE INTRODUCTION FOR MASS SPECTROMETRIC ANALYSIS

Richard N. Ellson and Mitchell W. Mutz, assignors to Picoliter Incorporated
5 August 2003 (Class 250/288); filed 14 February 2001

The salient portion of the devices described in this patent consists of an ejector, which consists of an acoustic radiation generator and a means for focusing the acoustic energy at a point near the surface of a small fluid sample.—EEU

6,603,241

43.35.Zc ACOUSTIC MIRROR MATERIALS FOR ACOUSTIC DEVICES

Bradley Paul Barber *et al.*, assignors to Agere Systems, Incorporated
5 August 2003 (Class 310/335); filed 23 May 2000

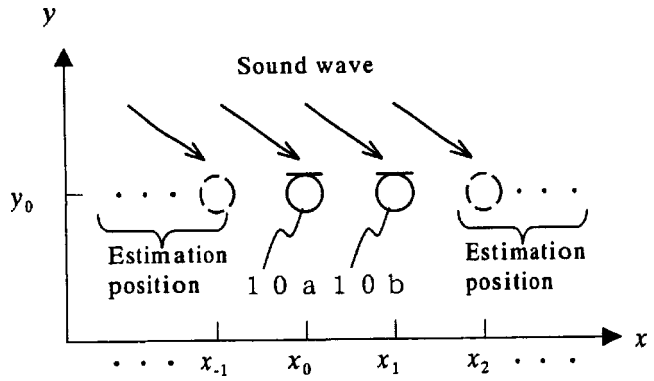
This patent is directed at bulk acoustic wave devices, primarily as used in cellular, wireless, and fiber-optic communications. Acoustic wave devices typically can be made much smaller than corresponding devices based on electromagnetic waves. An acoustic mirror according to this patent in essence consists of a substrate which supports alternating layers of materials with high and low acoustic impedance. The impedance mismatches desired here are in the range of about 10:1 to 6000:1 and are obtained by use of specific materials cited in the patent.—EEU

6,600,824

43.38.Hz MICROPHONE ARRAY SYSTEM

Naoshi Matsuo, assignor to Fujitsu Limited
29 July 2003 (Class 381/92); filed in Japan 3 August 1999

Some readers, including this reviewer, will find the explanation of this invention hard to follow. Suppose we have four people seated across the width of a table in a teleconferencing studio. In this situation, any talker can be identified in terms of the lateral angle from the camera to the source. The distance is assumed to be long enough that a plane wave impinges on microphones at locations **X0** and **X1** near the camera. One might measure the relative phase of the two microphone signals and use this information to steer the array. Instead, the patent goes through a lengthy mathematical analysis of sound pressure and velocity to compute the signals that would be



received at additional locations **X−1** and **X2**. The proposed procedure is intended to create a phantom array that is more directional than the actual array through electronic processing. Acoustical bootstrapping?—GLA

6,546,106

43.38.Ja ACOUSTIC DEVICE

Henry Azima, assignor to New Transducers Limited
8 April 2003 (Class 381/152); filed in the United Kingdom 2 April 1998

A flat loudspeaker with favorable frequency characteristics is formed from a panel with suitably selected edge restraints and distributed ribs that divide the panel into subpanels of different areas, thicknesses, and damping.—EEU

6,592,375

43.38.Ja METHOD AND SYSTEM FOR PRODUCING ENGINE SOUNDS OF A SIMULATED VEHICLE

Michael L. Henry *et al.*, assignors to Midway Games West Incorporated
15 July 2003 (Class 434/62); filed 9 February 2001

In a computer controlled arcade driving game, the gamer must be supplied with ample realistic exhaust sounds. The solution, say this patent's authors, is to place a loudspeaker at the end of a tube located underneath the seat. Since the authors believe that the tube ''focuses the sound much like a real engine exhaust pipe,'' it belies a fundamental misunderstanding of the nature of automobile acoustics. For what the authors really want, they would do better to use a big power amplifier and even larger loudspeakers.—MK

6,597,794

43.38.Ja PORTABLE ELECTRONIC DEVICE HAVING AN EXTERNAL SPEAKER CHAMBER

James R. Cole and Andrew L. Van Brocklin, assignors to Hewlett–Packard Development Company, L.P.
22 July 2003 (Class 381/333); filed 23 January 2001

A small portable computer usually sits on little feet, partly to provide ventilation to the underside of its case. The backwave from the device's loudspeaker might also be vented into this space. The invention adds some little ridges (acoustical dams) to enhance low-frequency energy. Well, if the computer is set on a perfectly flat, nonabsorptive surface that somehow makes an airtight seal with the acoustical dams, it just might work.—GLA

6,597,795

43.38.Ja DEVICE TO IMPROVE LOUDSPEAKER ENCLOSURE DUCT

Stephen Swenson, assignor to Stephen Swenson
22 July 2003 (Class 381/349); filed 25 November 1998

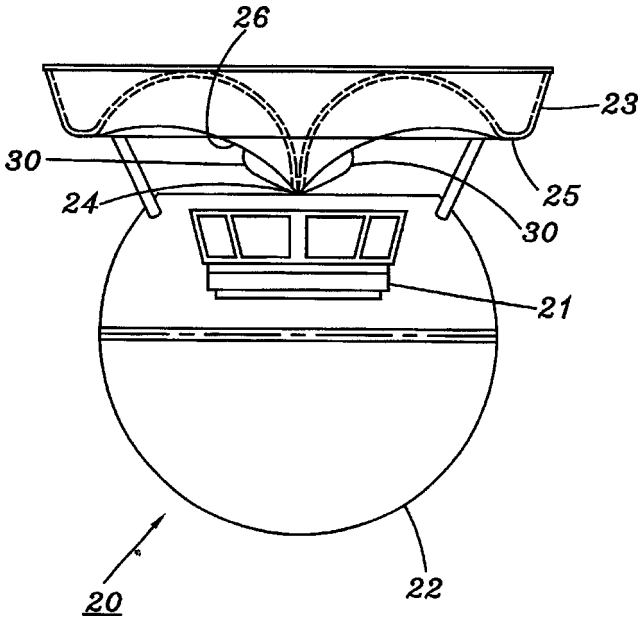
In a vented loudspeaker box, the vent tunnel is slotted or perforated, ''...increasing the duct impedance of the loudspeaker which improves transmission of the low-frequency waves to the exterior of the loudspeaker enclosure.'' In the real world such slots introduce losses that can substantially attenuate vent output.—GLA

6,597,797

43.38.Ja SPHERICAL LOUDSPEAKER SYSTEM WITH ENHANCED PERFORMANCE

Robert W. Betts, assignor to Sonic Systems, Incorporated
22 July 2003 (Class 381/366); filed 20 June 2000

This is complementary to United States Patent 6,603,862, also reviewed in this issue. Upward firing loudspeaker **21** faces reflector **23**. The reflector has a special sculpted shape, a little like an orange juicer, intended



to provide uniform hemispherical distribution of high-frequency energy. The patent assures us that this goal has been achieved but no test results are shown.—GLA

6,600,400

43.38.Ja ELECTROMAGNETIC ELECTRO-ACOUSTIC TRANSDUCER

Kimihiro Andou *et al.*, assignors to Matsushita Electric Industrial Company, Limited
29 July 2003 (Class 335/252); filed in Japan 7 September 1999

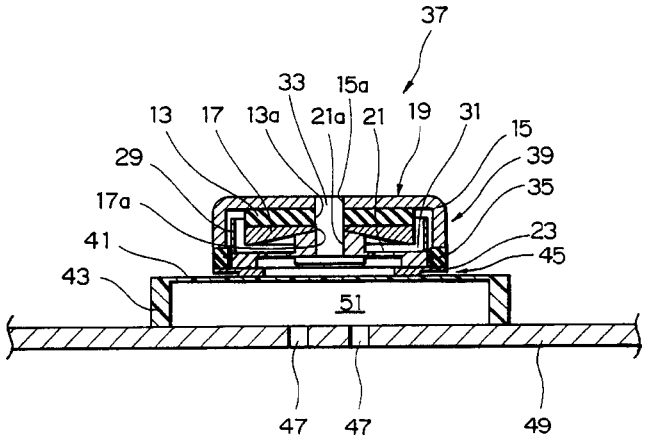
A simple electromagnetic transducer as used in cellular phones and pagers consists of a diaphragm, a magnet, and a coil. One might wonder what could be achieved with further complications. The patent explains in detail that a number of improvements can be realized if the ceramic magnet is formed from a combination of a hard magnetic material and a soft magnetic material.—GLA

6,600,938

43.38.Ja VIBRATION ACTUATOR AND MOBILE COMMUNICATION TERMINAL

Fumihiro Suzuki, assignor to NEC Tokin Corporation
29 July 2003 (Class 455/567); filed in Japan 17 May 1999

In a cellular phone or pager, inertia transducer **39** is normally mounted directly to case wall **49**, generating audible vibrations in the manner of a tuning fork. This patent suggests that greater acoustic output can be obtained



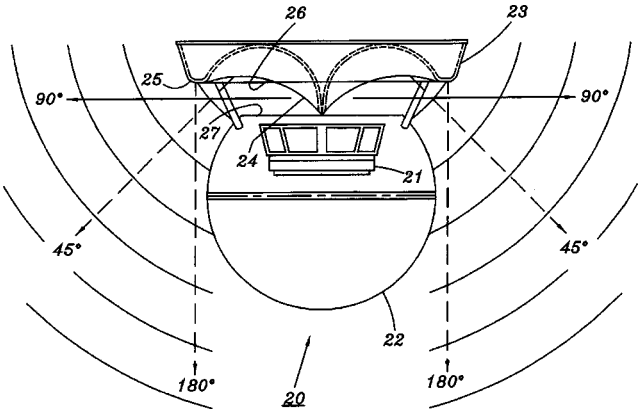
by mounting the vibrator to plate **41** which in turn drives Helmholtz resonator **51**. A double-chamber version is included in the six embodiments described.—GLA

6,603,862

43.38.Ja SPHERICAL LOUDSPEAKER SYSTEM

Robert W. Betts, assignor to Sonic Systems, Incorporated
5 August 2003 (Class 381/336); filed 25 October 1999

This patent complements United States Patent 6,597,797, reviewed above. The general configuration shown has enjoyed commercial success for a number of years. This latest patent describes the shape of reflector **23** in detail and discloses a mathematical formula that defines its contour. The



combined geometry of the reflector and spherical enclosure **22** is said to result in an optimum combination of horn loading, reflection, and diffraction. ''This ideal condition is known as maintaining a linear coupling coefficient through exponential expansion.''—GLA

6,603,863

43.38.Ja HEADPHONE APPARATUS FOR PROVIDING DYNAMIC SOUND WITH VIBRATIONS AND METHOD THEREFOR

Atsushi Nagayoshi, assignor to Matsushita Electric Industrial Company, Limited
5 August 2003 (Class 381/380); filed in Japan 25 December 1998

Old-fashioned headphones were supported by a headband on top of the user's head, like a hat. Newer lightweight units are mounted to a yolk that hangs beneath the user's chin. Suppose that the yolk were positioned behind the user's head instead, and suppose that it housed a vibration generator that rested on the user's neck. Mmmm, feels good.—GLA

6,604,602

43.38.Ja SEPARABLE SPEAKER COVER BOX CONTAINING SPEAKER SYSTEM

Chae Yong Kim, Ontario, California
12 August 2003 (Class 181/199); filed 30 September 2002

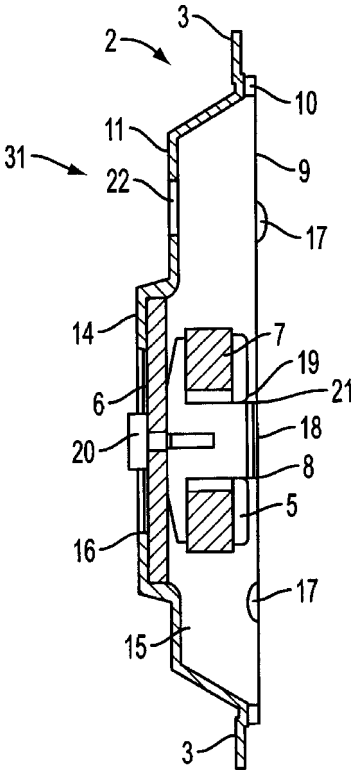
A conventional, wide-range speaker is fine for listening to classical music, but suppose you also want to listen to bass-pounding rock and roll? To meet both requirements this invention provides a detachable front chamber that converts the wide-range speaker into a tuned bandpass system.—GLA

6,606,390

43.38.Ja LOUDSPEAKERS

Henry Azima, assignor to New Transducer Limited
12 August 2003 (Class 381/152); filed in the United Kingdom 13 October 1998

The invention is a unitized panel form loudspeaker intended to reproduce mid and high frequencies. Enclosure **2** supports electrodynamic exciter **5** which drives panel **9** to generate bending waves. Small additional masses



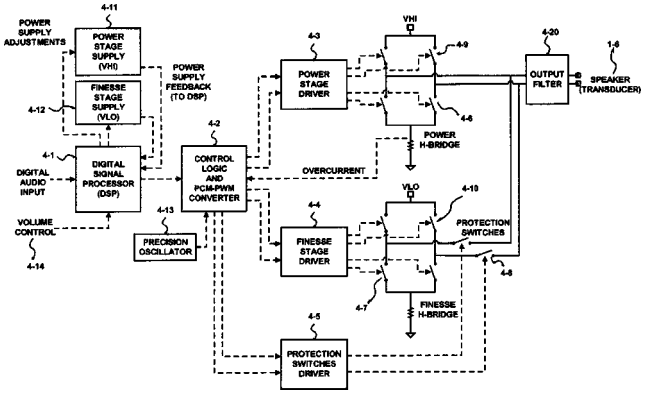
17 may be located to control unwanted resonances. The patent document includes a frequency response curve that shows relatively smooth response from about 250 Hz to beyond 20 kHz.—GLA

6,593,807

43.38.Lc DIGITAL AMPLIFIER WITH IMPROVED PERFORMANCE

William Harris Groves, Jr. and Arthur Alan Aaron, both of Arlington, Massachusetts
15 July 2003 (Class 330/10); filed 10 December 2001

Class D switching amplifiers are finding steadily increasing commercial applications, particularly in powered loudspeaker systems. This patent includes a concise, clearly written summary of the conflicts between a practical clock speed and high-quality performance. The circuit shown splits a digital audio signal into packets of higher and lower significant bits which



are then processed separately. The two output stages operate at the same clock frequency, but are time-division multiplexed. The combined signal is summed and averaged through a conventional reconstruction filter **20** before interfacing with the loudspeaker.—GLA

6,591,526

43.38.Md AUDIO-CAPABLE PHOTO ALBUM

Charles A. Garrett, Six Mile, South Carolina
15 July 2003 (Class 40/455); filed 18 June 2002

Basically, put a DSP microprocessor in the spine of a 3-ring photo binder. Add a microphone and speaker and you can now annotate a collection of photographs. The author doesn't address how to synchronize the pictures with the annotations or the explosion of digital photography—the latter renders this invention obsolete from the start.—MK

6,597,928

43.38.Md PORTABLE TELEPHONE MARACA WHICH PRODUCES MUSICAL TONES IN SYNCHRONIZATION OF THE SWINGS AND VIBRATION OF THE USER

Shuhei Ito, assignor to Yamaha Corporation
22 July 2003 (Class 455/567); filed in Japan 25 January 2000

Your cellular telephone is really a maraca. So says the inventor at Yamaha. All you need to add is a vibration sensor and a sound synthesis algorithm and you'll be ready to salsa.—MK

6,600,700

43.38.Md DIGITAL AUDIO DISC RECORDER

Kazuhisa Machida *et al.*, assignors to Denon, Limited
29 July 2003 (Class 369/30.19); filed in Japan 16 November 1999

On occasion, a patent is filed to protect a specific embodiment as opposed to an idea or concept. Denon's patent sketches the hardware and

software architecture of a digital audio workstation designed specifically for broadcast use. The novelty is nonexistant and the prior art search was performed only by the examiner.—MK

6,605,765

43.38.Md ACOUSTIC GUITAR WITH INTERNALLY LOCATED CASSETTE TAPE PLAYER

William A. Johnson, Arab, Alabama
12 August 2003 (Class 84/267); filed 18 January 2002

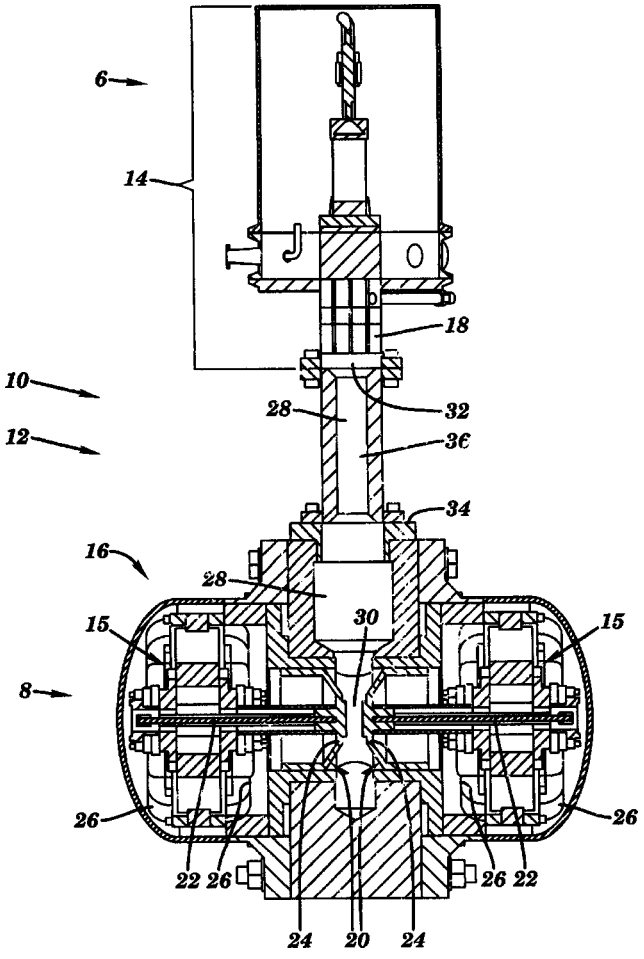
This patent both overstates and underspecifies the obvious. Clearly you could put a cassette tape recorder and player inside the body of a guitar—but in this day when tapes of all kinds are a vanishing species, this makes the patent already look antiquated.—MK

6,604,363

43.38.Pf MATCHING AN ACOUSTIC DRIVER TO AN ACOUSTIC LOAD IN AN ACOUSTIC RESONANT SYSTEM

John A. Corey and Jerry Martin, assignors to Clever Fellows Innovation Consortium; Mesoscopic Devices
12 August 2003 (Class 62/6); filed 19 April 2002

The invention has to do with pulse-type cryocoolers and the like which operate at a fixed resonant frequency. If such a system combines independently designed drivers and loads, it is unlikely that all conditions for optimum operation will be met. The patent explains how a properly sized matching volume **28** can be inserted between driver **8** and load **6** such that



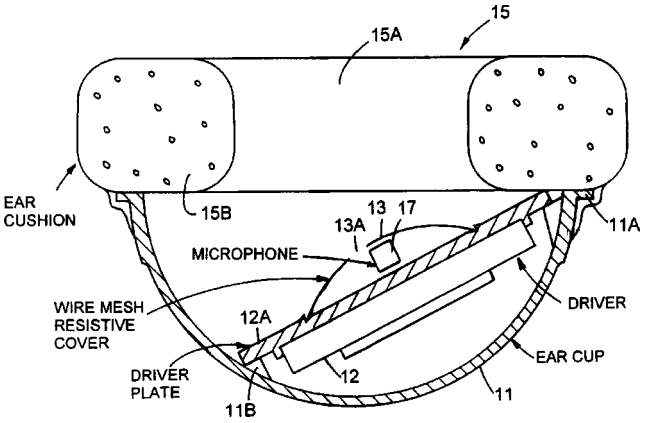
"...a resulting pressure wave delivers the preferred input acoustic flow amplitude to the load when the acoustic driver is operating approximately at the operating resonant frequency, the preferred stroke, and the preferred stroke amplitude."—GLA

6,597,792

43.38.Si HEADSET NOISE REDUCING

Roman Sapiejewski and Michael J. Monahan, assignors to Bose Corporation
22 July 2003 (Class 381/71.6); filed 15 July 1999

The illustration is a cross section through one earpiece of a Bose noise-cancelling headset. Two novel features are disclosed in this short patent. First, perforations in the skin of ear cushion **15** effectively increase



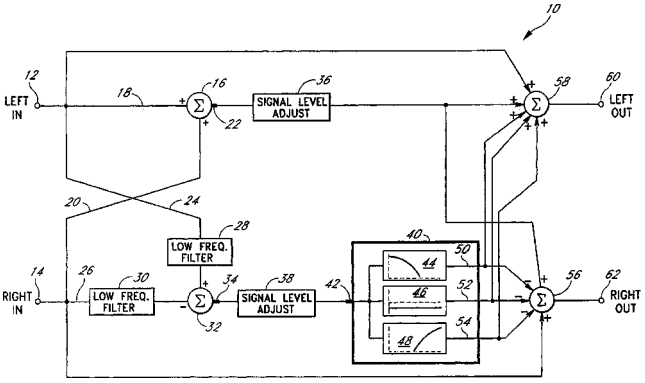
the internal volume of earcup **11**. Second, wire mesh screen **13** acts as an acoustic load that reduces the effects of resonances inside the earcup.—GLA

6,597,791

43.38.Vk AUDIO ENHANCEMENT SYSTEM

Arnold I. Klayman, assignor to SRS Labs, Incorporated
22 July 2003 (Class 381/1); filed 15 December 1998

In a two-channel stereo system, the apparent source width can be increased by increasing the relative level of the difference signal. This effect has long been known and forms the basis for many stereo enhancement circuits. The inventor's earlier patents disclose "critically acclaimed" techniques for filtering sum and difference signals in relation to human auditory characteristics. This latest "award winning" update is mainly concerned



with computer gaming applications that require small, simple, inexpensive circuitry. High-pass filters **28** and **30** attenuate low-frequency components in the difference signal. Equalizer **40** consists of separate low-pass, high-pass, and attenuation sections. Their combined effect produces a broad bump at 125 Hz followed by a dishout centered near 2 kHz and then a steadily rising characteristic that tapers off above 10 kHz.—GLA

6,591,949

43.40.Jc IMPACT ENERGY ABSORBING MEMBER

Akihiko Kitano *et al.*, assignors to Toray Industries, Incorporated
15 July 2003 (Class 188/371); filed in Japan 17 December 1999

The energy of an impact, such as occurs when vehicles collide, is absorbed by relatively light members made of fiber-reinforced resin. Energy absorbing members may be bonded or fixed mechanically to portions of a vehicle where impact protection is desired.—EEU

6,591,681

43.40.Le NONDESTRUCTIVE INSPECTION APPARATUS FOR INSPECTING AN INTERNAL DEFECT IN AN OBJECT

Takashi Shimada *et al.*, assignors to Mitsubishi Denki Kabushiki Kaisha
15 July 2003 (Class 73/600); filed 22 April 2002

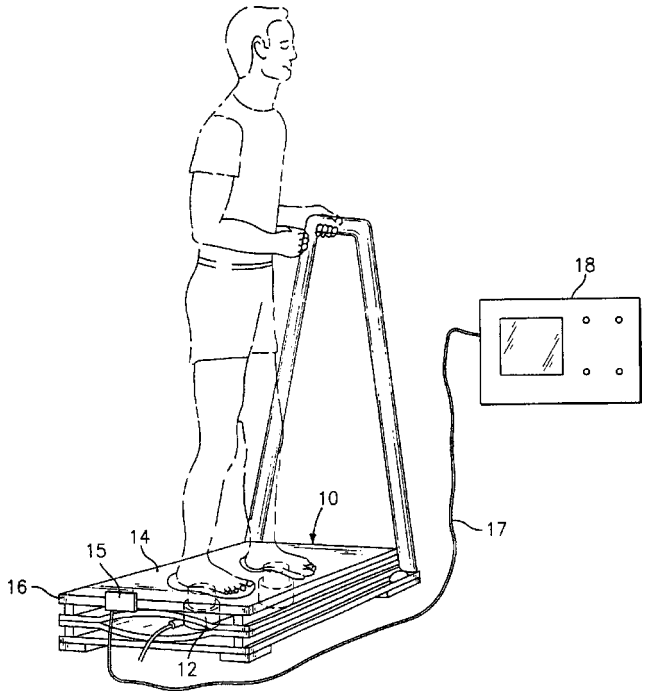
A magnetostrictive driver and a vibration sensor are pushed against the surface of a concrete element that is to be inspected. The apparatus includes means for providing the pushing forces and for measuring these. Its data processing accounts for the effects of these forces. The vibrator produces a ''chirp'' wave with continuously time-varying frequency, and the apparatus' signal processing section determines the distance to an internal defect by detecting the envelope of the reflected signal and using previously stored calibration data.—EEU

6,607,497

43.40.Ng NON-INVASIVE METHOD FOR TREATING POSTURAL INSTABILITY

Kenneth J. McLeod and Clinton T. Rubin, assignors to The Research Foundation of the State University of New York (SUNY)
19 August 2003 (Class 600/595); filed 18 December 2000

This patent, essentially a follow up of United States Patent 6,234,975 [reviewed in J. Acoust. Soc. Am. **110**(6), 2830 (2001)], discloses a method for treating postural instability due to, for example, the onset of senile sarcopenia. Diagnosis under the earlier patent is established by noninvasively measuring and analyzing the vibrational characteristics of the musculoskeletal system. The present patent seeks to provide a noninvasive method of treating the patient. This entails (a) providing a vibration table having a nonrigidly supported platform, (b) having a patient rest on the platform for a predetermined period of time, and (c) repeating steps (a) and (b) over a predetermined treatment duration. In step (b), a vibrational response of the patient's musculoskeletal system is measured and analyzed into specific vibrational spectra in order to evaluate the postural stability. The predetermined period of exposure is approximately ten minutes and the treatment duration is said to be at least four weeks.—DRR



6,547,049

43.40.Tm PARTICLE VIBRATION DAMPER

Geoffrey R. Tomlinson, assignor to Rolls–Royce plc
15 April 2003 (Class 188/379); filed in the United Kingdom 5 November 1999

One or more chambers are partially filled with many small particles. When this damper is attached to a vibrating object the particles impact against each other and against the chamber wall, thus removing energy from the primary vibration.—EEU

6,547,294

43.40.Tm SHOCK ABSORBING CONSTRUCTION FOR MOTORCYCLES

Kazuhiro Yamamoto, assignor to Honda Giken Kogyo Kabushiki Kaisha
15 April 2003 (Class 293/105); filed in Japan 27 December 2000

A shock absorber intended to cushion front or rear impacts on a motorcycle consists of a boxlike attachment that is made of plastic and filled with foam. A grillage of ribs within the plastic box provides strength and controlled collapse.—EEU

6,547,526

43.40.Tm ARTICLE HAVING DAMPENING MEMBER INSTALLED INTO AN IMBEDDED CAVITY

Thomas J. Van Daam and Timothy J. Hosking, assignors to The Boeing Company
15 April 2003 (Class 416/190); filed 3 July 2001

This patent pertains primarily to the damping of turbine blades. One or more circumferential cavities extend through the tips of all of the turbine blades and contain loosely fitting rings. Friction between these rings and the cavity surfaces damps the blade vibrations.—EEU

6,550,440

43.40.Tm ACOUSTIC SUPPRESSION ARRANGEMENT FOR A COMPONENT UNDERGOING INDUCED VIBRATION

Gary Allan Vrsek *et al.*, assignors to Ford Global Technologies, LLC
22 April 2003 (Class 123/184.61); filed 17 January 2002

Automotive engine components, such as intake manifolds made from polymeric materials, are friction-damped via elements that are snap-fit onto them.—EEU

6,550,868

43.40.Tm DAMPER PLATE FOR AUTOMOBILE WHEEL

Yuichi Kobayashi and Takumi Sekiguchi, assignors to The Yokohama Rubber Company, Limited
22 April 2003 (Class 301/6.91); filed in Japan 12 December 2000

A disc made of a lamination of steel and elastomeric plates is inserted between an automobile wheel and its axle in order to improve ride comfort and reduce interior noise.—EEU

6,551,057

43.40.Tm DAMPED TORQUE SHAFT ASSEMBLY

Frederic G. Haaser *et al.*, assignors to General Electric Company
22 April 2003 (Class 415/119); filed 22 November 1999

The torque shaft assembly to which this patent pertains is a shaft that is used to adjust the stator vanes on gas turbine engines. This shaft is not part of the rotating assembly, but actuates the stator vans via connecting elements. In order to limit the vibrations of the torque shaft, it is made in the form of a hollow tube that is filled with a granular material. A 98%-by-volume filling of steel shot is preferred.—EEU

6,554,112

43.40.Tm VIBRATION-DAMPING DEVICE FOR VEHICLE

Rentaro Kato *et al.*, assignors to Tokai Rubber Industries, Limited
29 April 2003 (Class 188/379); filed in Japan 8 September 2000

The devices described here in essence consist of metal cups in which metal balls can rattle when the assembly is subjected to vibrations. The balls are coated with rubber, as are some of the interior surfaces of the cups. On some cup surfaces, the rubber elements are supported only at their edges so that they tend to act as relatively soft springs, resulting in low natural frequencies for the affected ball rattling.—EEU

6,565,061

43.40.Tm RADIAL SNUBBER FOR VIBRATION ISOLATOR

Mark Edward Petersen and William P. Adkins, Jr., assignors to Honeywell International Incorporated
20 May 2003 (Class 248/638); filed 28 September 2000

In an isolator that is intended to act primarily in the radial direction, an inner cylindrical metal element is connected to an outer ring-shaped element via an elastomeric element in the shape of a conical shell. Snubbing of radial motions is provided by an added disc-shaped elastomeric element connecting the inner element to the outer element.—EEU

6,565,446

43.40.Tm ELASTIC SHAFT JOINT

Shoji Oka and Takahiro Minamigata, assignors to NSK Limited
20 May 2003 (Class 464/89); filed in Japan 11 June 1998

This patent pertains to a resilient element that is to be included in a universal joint in order to limit the transmission of axial vibrations. The design is said to be simple and inexpensive.—EEU

6,566,774

43.40.Tm VIBRATION DAMPING SYSTEM FOR ESP MOTOR

Larry James Parmeter and Dick L. Knox, assignors to Baker Hughes Incorporated
20 May 2003 (Class 310/90); filed 9 March 2001

This patent describes fluid film bearings for use on electric submersible pumps with vertical shafts. Lubricating fluid is supplied under pressure to lubricate and replenish the fluid film between the inner, rotating bearing sleeve and the outer, stationary sleeve insert.—EEU

6,567,212

43.40.Tm VIBRATION DAMPING DEVICE FOR MICROSCOPES AND MICROSCOPE WITH A VIBRATION DAMPING DEVICE

Johann Engelhardt and Rafael Storz, assignors to Leica Microsystems Heidelberg GmbH
20 May 2003 (Class 359/368); filed 16 August 2000

The device described here consists in essence of a mass that is connected via foam rubber to an optical microscope's structural element whose vibration is to be reduced. Typical structural elements are the sample stage of an upright microscope or the resolver of a confocal microscope. In many cases the vibrations of concern are those produced by motors that actuate the stages or focusing devices.—EEU

6,567,265

43.40.Tm APPARATUS HAVING FLEXIBLE MOUNTING MECHANISM

Toshiki Yamamura *et al.*, assignors to Matsushita Electric Industrial Company, Limited
20 May 2003 (Class 361/685); filed in Japan 20 November 1995

A portable recording or playback device that needs to be protected against vibration is placed into a housing and isolated from that housing. This housing, in turn, is inserted into a second housing, again with interposed vibration isolation. The two isolation systems are designed to have different characteristics.—EEU

6,568,666

43.40.Tm METHOD FOR PROVIDING HIGH VERTICAL DAMPING TO PNEUMATIC ISOLATORS DURING LARGE AMPLITUDE DISTURBANCES OF ISOLATED PAYLOAD

Worthington B. Houghton, Jr., assignor to Newport Corporation
27 May 2003 (Class 267/136); filed 13 June 2001

This patent describes an advanced design of an air-spring isolator or ''air leg.'' Damping is achieved by means of suitable gaps between coupled air volumes.—EEU

6,569,509

43.40.Tm ULTRALIGHT, SOUND AND SHOCK ABSORBING COMPONENT SET

Thorsten Alts, assignor to Rieter Automotive (International) AG
27 May 2003 (Class 428/73); filed in Switzerland 9 January 1998

This patent describes a panel configuration intended for use in automobile interiors. A typical panel consists of an acoustically absorptive covering layer atop a crushable layer, such as honeycomb, mounted on a base layer. The crushable layer is intended to absorb shocks in impacts; adjacent air spaces in that layer are interconnected to enhance the acoustic absorption performance of the assembly.—EEU

6,575,037

43.40.Tm MULTIPLE DEGREE OF FREEDOM VIBRATION EXCITING APPARATUS AND SYSTEM

Yasuyuki Momoi *et al.*, assignors to Hitachi, Limited
10 June 2003 (Class 73/633); filed 25 July 2001

This patent very obscurely describes an apparatus for exciting a test structure in two translational degrees and one rotational degree of freedom in a plane. The apparatus consists of three actuators and collocated force and motion sensors plus a computational and control system. The system tests part of the structure, calculates the desired motion for an added part of the structure, and couples the results to determine the excitation of the combined structures.—EEU

6,578,836

43.40.Tm SHOCK ABSORBER FOR AUTOMOBILE

Tomohiko Kogure *et al.*, assignors to The Yokohama Rubber Company, Limited
17 June 2003 (Class 267/152); filed in Japan 14 December 2000

This shock absorber consists of a steel coil spring on the ends of which there are located plates of laminated elastic layers. These layers have their peak loss factors in different frequency ranges encompassing several of the spring's natural frequencies.—EEU

6,585,240

43.40.Tm VIBRATION RELIEF APPARATUS AND MAGNETIC DAMPER MECHANISM THEREFOR

Etsunori Fujita and Yasuhide Takata, assignors to Delta Tooling Company, Limited
1 July 2003 (Class 267/136); filed in Japan 19 August 1999

This patent appears to pertain primarily to seats in automobiles and other conveyances. It consists in essence of a mechanical linkage that uses springs and levers to obtain a very small effective stiffness in the vertical direction over a significant range of excursions. Damping is obtained by permanent magnet eddy current devices attached to parts that move relative to each other.—EEU

6,585,902

43.40.Tm NON-HOMOGENEOUS LAMINATE MATERIAL FOR SUSPENSION WITH VIBRATION DAMPENING

Victor Wing–Chun Shum and Randall George Simmons, assignors to International Business Machines Corporation
1 July 2003 (Class 216/13); filed 27 February 1999

This patent relates in general to integrated lead suspensions for hard disk drives, which are essentially in the form of cantilever beams. A lead suspension according to this patent is made from a five-layer laminate of stainless steel, polyimide, and copper. The steel layer has preformed voids in order to weaken the high-strain flexure areas, thereby increasing the energy dissipation capability of the suspension.—EEU

6,588,554

43.40.Tm VIBRATION DAMPING APPARATUS USING MAGNETIC CIRCUIT

Etsunori Fujita *et al.*, assignors to Delta Tooling Company, Limited
8 July 2003 (Class 188/267); filed in Japan 15 November 2000

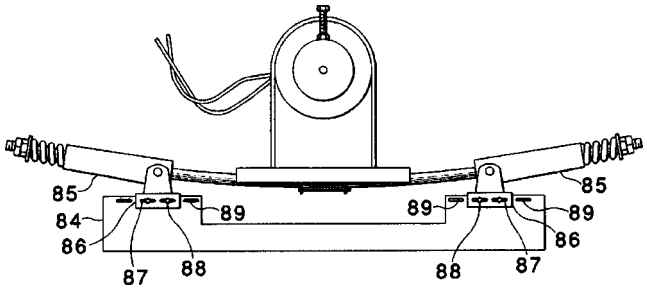
This vibration isolator is configured essentially like a piston that is relatively free to move axially in a cylinder. Motion in the axial direction is countered by an axial coil spring, with damping at small amplitudes provided by interacting rings of permanent magnets that are attached to the piston and cylinder portions. Damping at larger amplitudes is provided by a viscous fluid that is forced through a series of shaped passages between the circumference of the piston and the inside of the cylinder. The piston shaft is guided by an elastomeric bushing, which provides some resilience in the presence of nonaxial disturbances.—EEU

6,595,483

43.40.Tm VIBRATION ISOLATION DEVICE AND METHOD

John Cunningham, Saratoga Springs, New York
22 July 2003 (Class 248/609); filed 16 February 2001

Isolation is provided by a flexible member that is supported on pivots via low-friction sleeves **85**. The natural frequency of the arrangement may



be changed by adjusting the distance between pivots by making use of a series of slots **87**, **88**, and by modifying the tension in the axial springs.—EEU

6,598,545

43.40.Tm VIBRATION DAMPER FOR OPTICAL TABLES AND OTHER STRUCTURES

Vyacheslav M. Ryaboy and Worthington B. Houghton, Jr., assignors to Newport Corporation
29 July 2003 (Class 108/136); filed 10 April 2001

An optical table in essence consists of two parallel flat face plates, assembled so that they are at some distance from each other. The present patent describes a series of small plates that are mounted around the edges of the primary plates and perpendicular to them. Some of these small plates are attached only to the upper face plate, and some only to the lower face plate, with plates from the two sets near each other and interconnected via a layer of viscoelastic material. As the table flexes, the small plates tend to move relative to each other, so that the layer is deflected in shear, thus damping the flexural motion.—EEU

6,598,718

43.40.Tm VIBRATION DAMPING DEVICE HAVING A FLOCK COATING

Rentaro Kato and Koichi Maeda, assignors to Tokai Rubber Industries, Limited
29 July 2003 (Class 188/378); filed in Japan 30 March 2001

This patent pertains to dampers of the impact type, where a mass is essentially free to rattle within a somewhat oversized cavity. Here, one or more of the interacting surfaces is provided with a flocked coating, which absorbs some of the energy associated with the rattling impacts.—EEU

6,599,439

43.40.Tm DURABLE MAGNETORHEOLOGICAL FLUID COMPOSITIONS

Vardarajan R. Iyengar and Robert T. Foister, assignors to Delphi Technologies, Incorporated
29 July 2003 (Class 252/62.52); filed 14 December 2000

Magnetorheological fluids are suitable for use in controllable vibration attenuation devices. This patent describes long-term durable fluids comprised of mechanically hard magnetizable particles in a specific carrier fluid.—EEU

6,604,735

43.40.Tm ELASTOMER VARIANTS

Trevor J. McCollough and Robert James Monson, assignors to Lockheed Martin Corporation
12 August 2003 (Class 267/140); filed 29 November 2001

The title of this patent fails to indicate that the subject is a modular shock isolator, where a module essentially consists of a plate to which there is attached a relatively thin-walled elastomeric trapezoidal element. The latter tends to collapse, at least partially, in the presence of a load that acts perpendicular to the base plate. Modules can be used in various mechanical series and parallel arrangements.—EEU

6,590,639

43.40.Vn ACTIVE VIBRATION ISOLATION SYSTEM HAVING PRESSURE CONTROL

Bausan Yuan *et al.*, assignors to Nikon Corporation
8 July 2003 (Class 355/75); filed 15 September 2000

This patent relates particularly to isolation systems used in photolithography processing. The isolation arrangement described here consists of a combination of a pneumatic system and an electrodynamic system acting mechanically in parallel. The pneumatic system in effect is an air-spring, whose pressure is sensed and controlled via a processor in a feedback loop, combined with a feed-forward signal derived from the motion of the isolated mass. The electrodynamic system similarly uses an actuator that acts on the mass as directed by a controller, which receives feedback and feed-forward signals from mass motion and position sensors.—EEU

6,598,885

43.40.Vn SINGLE VALVE CONTROL OF DAMPING AND STIFFNESS IN A LIQUID SPRING SYSTEM

Damon Delorenzis *et al.*, assignors to Liquidspring Technologies, Incorporated
29 July 2003 (Class 280/5.507); filed 23 October 2001

Liquid springs, which in essence are liquid-filled cylinder/piston shock absorbers, can provide both flexibility and damping. Control of liquid flow to and from the spring's chambers permits adjustment of these parameters. The present patent describes an automotive suspension system that makes use of liquid springs, where the spring characteristics during jounce and rebound are controlled on the basis of signals from height, speed, and steering wheel position sensors.—EEU

6,600,987

43.40.Vn APPARATUS AND METHOD FOR DETERMINING A ROAD-WHEEL VIBRATION OF AUTOMOTIVE VEHICLE, AND APPARATUS AND METHOD FOR ANTI-SKID CONTROL USING THE SAME

Nobuyuki Ohtsu, assignor to Unisia Jecs Corporation
29 July 2003 (Class 701/71); filed in Japan 29 June 2000

In this anti-skid braking system, the pressure provided to each brake cylinder is controlled on the basis of measured oscillations of the wheel speeds. A sensor measures the speed of each road wheel and a calculating unit determines the accelerations and decelerations of each wheel during pressure application and release cycles. By comparing the measured vibrations with predetermined threshold values the unit determines whether the wheel vibration is caused by drive train vibration or by the road surface conditions and causes skid control to be actuated accordingly.—EEU

6,543,620

43.40.Yq SMART SCREENING MACHINE

Daryoush Allaei, assignor to Quality Research, Development & Consulting, Incorporated
8 April 2003 (Class 209/365.1); filed 23 February 2001

Distributed electrically controlled transducers are used to vibrate a screen used for separation of materials. This concept is said to result in screens that are smaller and less noisy than conventional ones. It also permits automatic detection of clogging and adjustment of the transducers to provide unclogging.—EEU

6,545,762

43.40.Yq METHOD OF INVESTIGATING VIBRATIONS AND AN APPARATUS THEREFOR

Harvey Lewis and Andrew Rogoyski, assignors to Logica UK Limited
8 April 2003 (Class 356/502); filed in the United Kingdom 17 April 2000

As in other laser vibrometers, light from a laser is split, with part being directed at the test object and part being mixed with light reflected from the object to create an interference signal. The latter is processed to derive a signal representing the object's vibrations. Unlike other laser vibrometers, which do not work in the presence of bulk motion of the test object, the system described in this patent derives an estimate of the frequency of the bulk movement from the observed signal and uses this frequency to reduce the bandwidth of the observation, so as to limit the interference from the bulk motion.—EEU

6,590,832

43.40.Yq VIBRATOR AND METHOD OF EXPLORING A MATERIAL MEDIUM BY MEANS OF VERY LOW-FREQUENCY ELASTIC VIBRATIONS

Jean-Claude Dubois and Marc Becquey, assignors to Institut Francais Du Petrole
8 July 2003 (Class 367/92); filed in France 1 February 2000

This patent pertains to seismic prospecting and similar applications. A sinusoidal force generator acts on the surface of the ground via at nonlinear spring element, resulting in injection into the ground of a signal with at least two frequency components. These frequencies can be varied within certain ranges by adjustment of the driving frequency. The sum and difference signals resulting from the injected components are sensed remotely and correlated with the driving signals in order to extract information concerning the intervening soil.—EEU

6,604,583

43.40.Yq VIBRATING DEVICE AND A METHOD FOR DRIVING AN OBJECT BY VIBRATION

Johan Bernard Van Randen, assignor to International Construction Equipment B.V.
12 August 2003 (Class 173/49); filed in the Netherlands 19 March 1998

This patent pertains to a vibrator for use in driving of piles or sheet piling. A vibrator according to this patent consists of two pairs of gear-connected, counter-rotating unbalanced weights, with the pairs interconnected via a differential ''phase shifter.'' The latter permits the phase of the motion of the two pairs to be changed, permitting adjustment of the net force amplitude provided by the device.—EEU

6,609,428

43.40.Yq NONRESONANT TECHNIQUE FOR ESTIMATION OF THE MECHANICAL PROPERTIES OF VISCOELASTIC MATERIALS

Andrew J. Hull, assignor to The United States of America as represented by the Secretary of the Navy
26 August 2003 (Class 73/789); filed 19 January 2001

This is a method for estimating the real and imaginary Young's modulus, shear modulus, and Poisson's ratio of a specimen at an excitation frequency. The specimen is first mounted onto a reciprocating test apparatus at one end, with a mass attached at the other end. The test apparatus reciprocates at the excitation frequency. Accelerations are recorded at each end of the specimen. The Young's modulus is computed from the accelerations. The specimen is then attached to a reciprocating rotational test apparatus at one end, with a rotational intertial mass positioned at the other end. Accelerations are recorded while subjecting the specimen to rotational reciprocations at the excitation frequency. The shear modulus is computed from these accelerations. Poisson's ratio can then be calculated from the Young's modulus and the shear modulus at the excitation frequency.—DRR

6,550,574

43.50.Gf ACOUSTIC LINER AND A FLUID PRESSURIZING DEVICE AND METHOD UTILIZING SAME

Zheji Liu, assignor to Dresser–Rand Company
22 April 2003 (Class 181/286); filed 21 December 2000

In order to obtain quieter centrifugal compressors, a disc-shaped liner consisting of an array of Helmholtz resonators is installed near the compressor's exit volute. This liner, which must be rugged, in essence consists of a plate into which there are drilled a number of large-diameter holes that do not extend entirely through its thickness, with a number of smaller holes drilled through the remaining thickness.—EEU

6,604,603

43.50.Gf SOUNDPROOFING FOR INSULATING SOUND PRODUCING DEVICES OR PARTS OF SYSTEMS, ESPECIALLY DEVICES THAT TRANSMIT VIBRATIONS SUCH AS VIBRATORS

Anton Wirth, assignor to Etis AG
12 August 2003 (Class 181/200); filed in the European Patent Office 17 December 1998

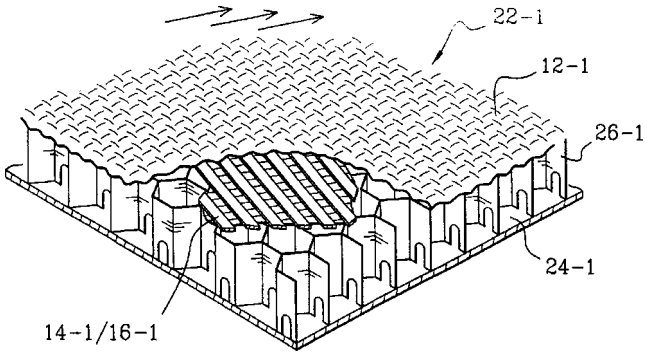
An acoustical enclosure according to this patent consists of relatively rigid walls, possibly made of metal or wood with a rubber layer between them, plus an interior liner. This liner consists of a relatively coarse innermost layer atop a layer of soft fiber mat material. The innermost layer is provided with multiple perforations so that sound can reach the fiber mat layer with the result that good absorption is obtained over an extended frequency range. The vibrating element housed in the enclosure is supported on a soft foam layer, thereby limiting structure-borne sound transmission. The absorptive layers on the interior walls are attached by velcro strips, thus avoiding the need for connections that could constitute sound-transmitting bridges.—EEU

6,607,625

43.50.Gf PROCESS FOR THE PRODUCTION OF AN ACOUSTIVELY RESISTIVE LAYER, RESISTIVE LAYER THUS OBTAINED, AND WALL USING SUCH LAYER

Robert Andre *et al.*, assignors to EADS Airbus SA
19 August 2003 (Class 156/169); filed in France 24 December 1999

This patent describes a process for producing a monolithic acoustical shield assembly, as for the nacelles of aircraft jet engines. The process starts with a structural component of thermoplastic resins that have an open surface and that are connected to an open metallic mesh. The openings in the



mesh are adapted to the open surface of the structural component. Polymerization of the resins under pressure and at high temperature bond the elements into one skin.—CJR

6,598,636

43.50.Lj METHOD FOR FIXING A RING MADE OF A SOUND-ABSORBING MATERIAL ON A RIM AND MOTOR VEHICLE WHEEL HAVING A RING MADE OF SOUND-ABSORBING MATERIAL FIXED ON THE RIM

Oliver Schürmann *et al.*, assignors to Continental Aktiengesellschaft
29 July 2003 (Class 152/450); filed in Germany 29 April 1998

This patent describes placement of a ring of sound-absorbing material on the rim of a tire, inside of the tire. A series of threads is used to fasten the ring to the rim, in order to keep it in place in the presence of the centrifugal forces.—EEU

6,543,719

43.50.Nm OSCILLATING AIR JETS FOR IMPLEMENTING BLADE VARIABLE TWIST, ENHANCING ENGINE AND BLADE EFFICIENCY, AND REDUCING DRAG, VIBRATION, DOWNLOAD AND IR SIGNATURE

Ahmed A. Hassan *et al.*, assignors to McDonnell Douglas Helicopter Company
8 April 2003 (Class 244/17.13); filed 24 November 1998

Porous surfaces are provided on aircraft surfaces, such as rotor blades, and are driven with oscillating pressures either continually or when predetermined flight conditions occur. These pressures control the aerodynamics of the surfaces and are said, among other things, to be able to produce significant reductions in the blade-vortex-interaction noise of helicopter rotors.—EEU

6,550,332

43.50.Yw REAL-TIME NOISE SOURCE VISUALIZING SYSTEM USING ACOUSTIC MIRROR

Myung-Han Lee, assignor to Hyundai Motor Company
22 April 2003 (Class 73/583); filed in the Republic of Korea 22 November 2000

This system, used in wind tunnel tests of automobiles, employs a microphone located at the focus of a parabolic reflector. The microphone system and a camera connected to it can be traversed to focus on selected areas. A display system provides color-coded sound images.—EEU

6,596,389

43.55.Ev FOAMED COMPOSITE PANEL WITH IMPROVED ACOUSTICS AND DURABILITY

Chester W. Hallett *et al.*, assignors to AWI Licensing Company
22 July 2003 (Class 428/317.9); filed 17 October 2000

This acoustical panel is formed with additives that induce voids evenly within the panel. This allows the panels to achieve very high sound absorption properties without the need for additional surface perforations. The panels also maintain a very high surface hardness. These additives are combined in a wet or dry form by the use of a high-intensity mixer that agitates and aerates the mixture to create an aerated slurry.—CJR

6,601,673

43.55.Ev SOUND ABSORBING STRUCTURE

Atsushi Murakami *et al.*, assignors to Nichias Corporation
5 August 2003 (Class 181/293); filed in Japan 6 September 2000

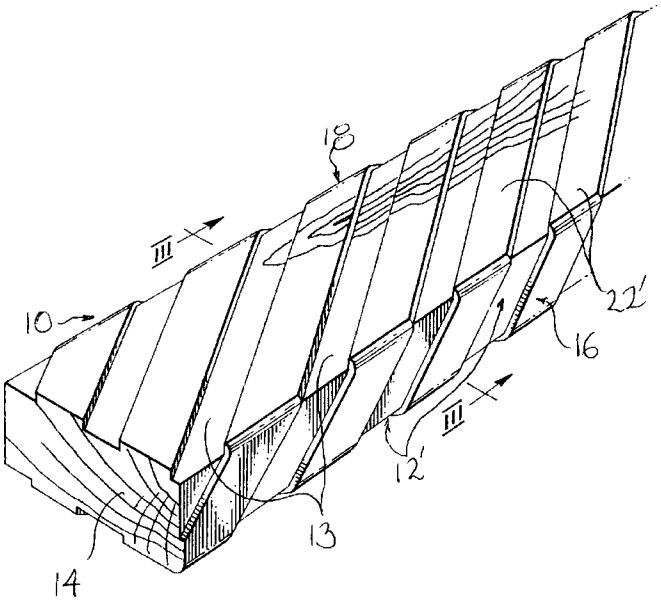
A finely perforated film is laminated onto a foam absorber. Then a perforated face is used to protect the foam. The size of the perforations of the barrier facing and the thickness of the foam determine the frequencies at which the absorber is most effective. The structure would be used to absorb sound from an automobile engine or the like.—CJR

6,594,964

43.55.Ti GROOVED CONSTRUCTION BEAM

Gilles Charland, St–Chrysostome, Quebec, Canada
22 July 2003 (Class 52/302.1); filed in the United Kingdom 15 August 2000

This so-called construction beam (really just a wood stud or joist) has grooves on some or all of its sides. These grooves reduce surface area of contact with wall sheathing or subfloor. This, in turn, reduces transmission



of sound and vibration and at the same time allows greater circulation of air to enhance thermal insulation and reduce trapped moisture.—CJR

6,604,715

43.55.Vj EXTRUDED ACOUSTIC PIPE SUPPORT

Samuel H. Howe, assignor to LSP Products Group, Incorporated
12 August 2003 (Class 248/62); filed 23 July 2001

These pipe supports or straps are for acoustical isolation and vibration damping. They have both a foamed cushion layer and an unfoamed skin layer that have been extruded together. The supports can be attached to structural support members using nails, screws, formed wire hangers, or other similarly effective fasteners.—CJR

6,599,195

43.58.Ta BACKGROUND SOUND SWITCHING APPARATUS, BACKGROUND-SOUND SWITCHING METHOD, READABLE RECORDING MEDIUM WITH RECORDING BACKGROUND-SOUND SWITCHING PROGRAM, AND VIDEO GAME APPARATUS

Shigeru Araki *et al.*, assignors to Konami Company, Limited; Kabushiki Kaisha Konami Computer Entertainment Osaka
29 July 2003 (Class 463/43); filed in Japan 8 October 1998

In previously reviewed United States Patent 6,540,613 [reviewed in J. Acoust. Soc. Am. **114**(3), 1208 (2003)], as the video game moves from scene to scene, the background sound changes frame synchronously. This patent addresses the issues of fading around frame boundaries. This borders on the (patently) obvious.—MK

6,591,524

43.58.Wc ADVERTISING ARTICLE WITH AUTOMATICALLY ACTIVATED FLASHER OR SOUND MODULE

Edward D. Lewis and Timothy D. Hogue, assignors to Buztronics, Incorporated
15 July 2003 (Class 40/324); filed 15 October 1996

In previous reviews (United States Patents 6,335,691 [reviewed in J. Acoust. Soc. Am. **112**(1), 23 (2002)] and 6,478,152 [reviewed in J. Acoust. Soc. Am. **113**(5), 2387 (2003)]), we have seen application of the ubiquitous sound chip to drinking vessels of all varieties. Here, the authors propose using a piezoelectric transducer to signal when the drink container is placed in its sleeve. Then lights and/or sound can appear, supplying much needed advertisement.—MK

6,592,260

43.58.Wc FLEXIBLE STORAGE BAG WITH AUDIBLE CLOSURE INDICATOR

Catherine Jean Randall *et al.*, assignors to The Procter & Gamble Company
15 July 2003 (Class 383/64); filed 18 August 2000

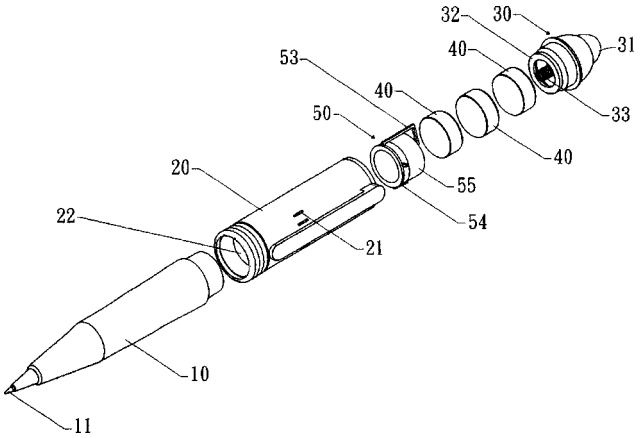
How do you know when a plastic bag is sealed? One manufacturer offers two strips of differing colors that blend when sealed. Here, Proctor and Gamble propose using a slider that creates pops or clicks as it seals.—MK

6,604,880

43.58.Wc MOTION LIGHTING PEN WITH LIGHT VARIABLY ACCOMPANYING SOUND ACTUATION

Fang-Yue Huang and Wen-Chen Chung, assignors to Excellence Optoelectronics, Incorporated
12 August 2003 (Class 401/195); filed 13 June 2002

In the continuing effort to make everything in the environment sound off, these inventors have added sound to a pen. As shown, transducer **55** is



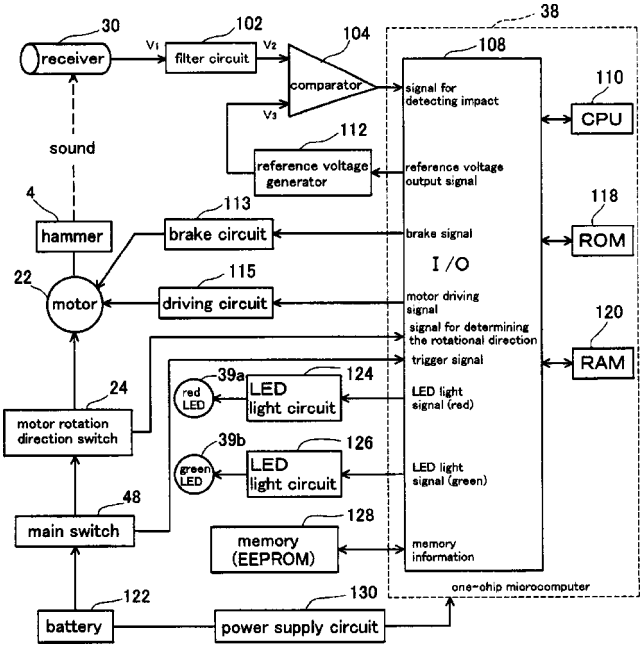
powered by batteries **40** and the principal advance, say the authors, is the use of a cylinder **20** as a resonant chamber (at what frequency?). Don't forget the LEDs!—MK

6,607,041

43.58.Wc POWER TOOLS

Hitoshi Suzuki and Masahiro Watanabe, assignors to Makita Corporation
19 August 2003 (Class 173/4); filed in Japan 16 March 2000

The days of simple, stupid power tools are over. Makita, a well-known Japanese manufacturer, proposes sensing of impact sounds by a microprocessor as shown in the figure. The impact detection method proposed (thresholding) is particularly simple minded.—MK



6,592,516

43.60.Bf INTERACTIVE CONTROL SYSTEM OF A SEXUAL DELIGHT APPLIANCE

Ching-Chuan Lee, Taipei, Taiwan, Province of China
15 July 2003 (Class 600/38); filed 9 October 2001

The abstract says it all: ''An interactive control system of a sexual delight appliance having a signal receiving circuit, a sound analyzing circuit, a status signal input circuit, and a mechanism driving circuit. The signal receiving circuit has an input port for receiving a foreign sound signal. The sound analyzing circuit analyzes and recognizes the received sound signal in terms of the volume, the tone, and the speed. The status signal input circuit converts the analyzed sound signal into a plurality of motion signals. The mechanism driving circuit performs the motion signals and allows at least one driving device to produce motions related to the motion signals such that a doll is moved with corresponding motions. In addition, the doll has a built-in sensing feedback device to remotely control another doll at a remote location via a communication net.''—MK

6,608,549

43.60.Lq VIRTUAL INTERFACE FOR CONFIGURING AN AUDIO AUGMENTATION SYSTEM

Elizabeth D. Mynatt *et al.*, assignors to Xerox Corporation
19 August 2003 (Class 340/5.8); filed 31 July 1998

A virtual interface is provided that allows a user to navigate through a representation of a physical target area, such as an office, a classroom, or home environment. The system transmits information to selected end users via peripheral or background auditory cues in response to physical actions of the end users in the environment. With the use of the virtual interface, the user can alter the system configuration. Examples of use are signals to indicate volume of e-mail, receipt of e-mail from a specific party such as a supervisor, the presence or absence of a worker, lapping activity of a water tank in a laboratory, etc.—DRR

6,567,006

43.60.Qv MONITORING VIBRATIONS IN A PIPELINE NETWORK

Paul Lander and William E. Saltzstein, assignors to Flow Metrix, Incorporated
20 May 2003 (Class 340/605); filed 19 November 1999

This system for detection of leaks in pipelines uses a series of ''locally intelligent, data-adaptive'' monitoring systems that are either synchronized or connected digitally to a base station. Received vibration signals from two or more systems are correlated to determine the position of any leak present in the pipeline.—EEU

6,568,271

43.60.Qv GUIDED ACOUSTIC WAVE SENSOR FOR PIPELINE BUILD-UP MONITORING AND CHARACTERIZATION

Vimal V. Shah *et al.*, assignors to Halliburton Energy Services, Incorporated
27 May 2003 (Class 73/599); filed 8 May 2001

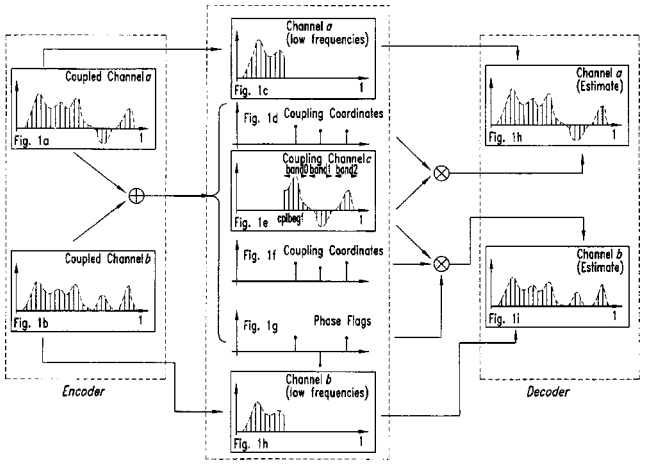
Deposits or buildup inside a fluid-containing pipe are monitored with the aid of a pair of acoustic transmitters that are mounted outside the pipe at some distance apart. A pair of acoustic receivers is mounted between the transmitters, spaced apart along the length of the pipe. The transmitters are configured to induce Lamb waves in the pipe wall and are actuated one at a time and the resulting amplitudes at the receivers are observed. The characteristics of the pipe wall (and whatever adheres to it) are determined from the measured amplitudes.—EEU

6,574,602

43.72.Gy DUAL CHANNEL PHASE FLAG DETERMINATION FOR COUPLING BANDS IN A TRANSFORM CODER FOR HIGH QUALITY AUDIO

Mohammed Javed Absar *et al.*, assignors to STMicroelectronics Asia Pacific PTE Limited
3 June 2003 (Class 704/504); filed 8 September 2000

Described is a method of coupling channels, based on a psychoacoustic model of the auditory system, for obtaining data compression in a dual channel audio encoder. Subband phase flags are determined to avoid phase



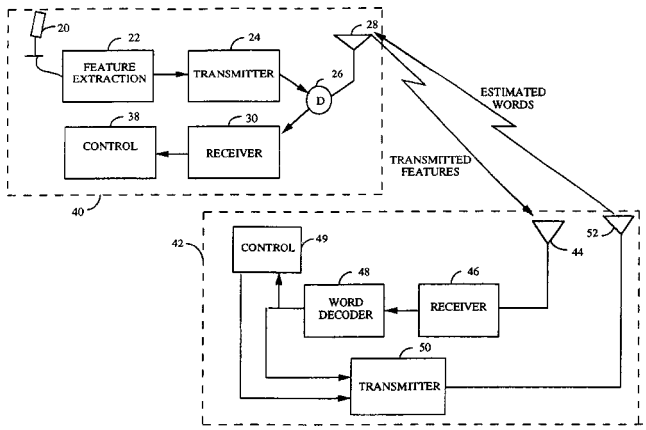
cancellation when the bands are combined. Least square error is ensured between the original channel frequency coefficients at the encoder and the estimated coefficients at the decoder.—DAP

6,594,628

43.72.Ne DISTRIBUTED VOICE RECOGNITION SYSTEM

Paul E. Jacobs and Chienchung Chang, assignors to Qualcomm, Incorporated
15 July 2003 (Class 704/231); filed 2 April 1997

A feature extraction mechanism for a distributed voice recognition system is located in a remote device, which, by separation from the word decoder, is said to provide improved system performance. The extracted features are provided to a word decoder in the central processing station that



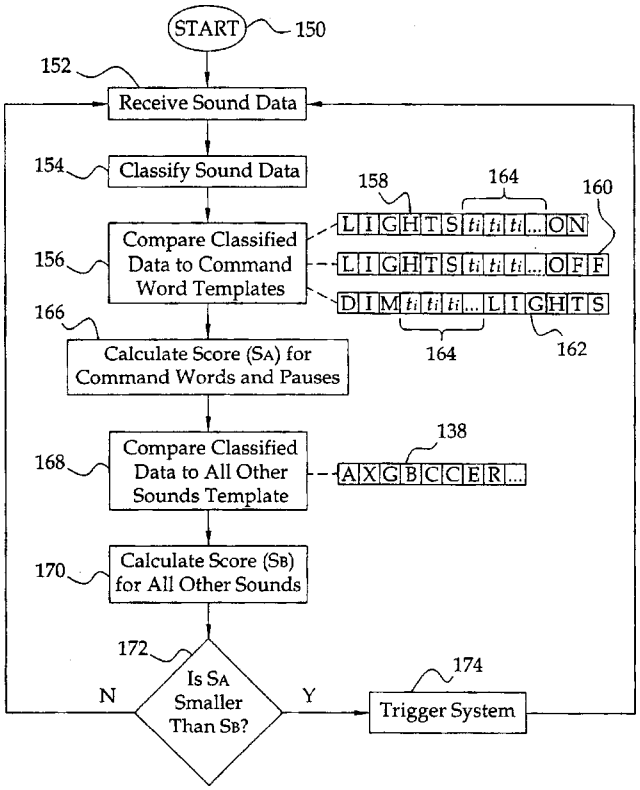
determines the original word string in the input speech frames. The front end processor is either LPC or filter-band based and the acoustic pattern matching in the word decoder is based on hidden Markov models, dynamic time warping, or neural networks.—DAP

6,594,630

43.72.Ne VOICE-ACTIVATED CONTROL FOR ELECTRICAL DEVICE

Igor Zlokarnik and Daniel Lawrence Roth, assignors to Voice Signal Technologies, Incorporated
15 July 2003 (Class 704/256); filed 19 November 1999

A voice-activated unit controls an electrical device.—HHN



6,598,017

43.72.Ne METHOD AND APPARATUS FOR RECOGNIZING SPEECH INFORMATION BASED ON PREDICTION

Naoko Yamamoto *et al.*, assignors to Canon Kabushiki Kaisha
22 July 2003 (Class 704/251); filed in Japan 27 July 1998

A module predicts the type of sound information using a knowledge-based expert system for speech recognition.—HHN

6,606,280

43.72.Ne VOICE-OPERATED REMOTE CONTROL

Guenter Knittel, assignor to Hewlett–Packard Development Company
12 August 2003 (Class 367/198); filed 22 February 1999

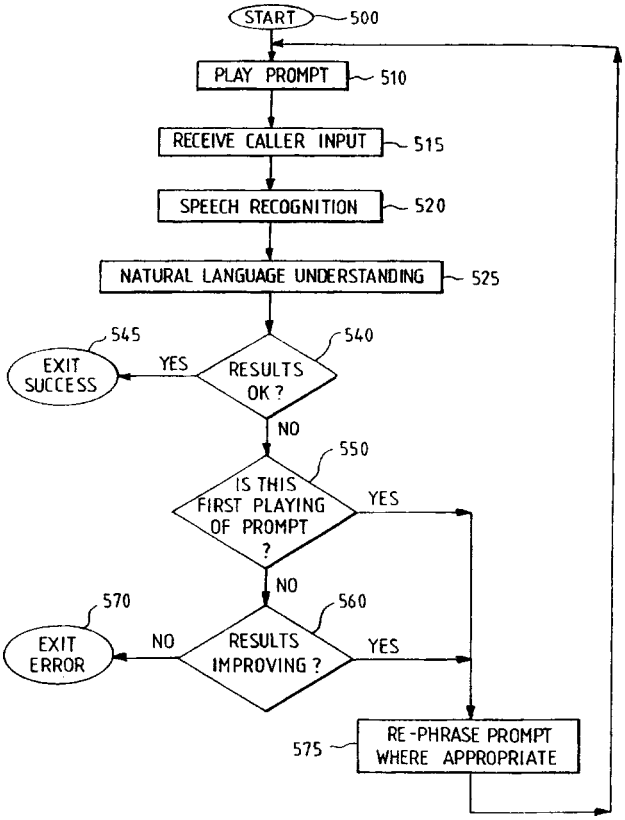
Why aren't remote controls voice activated? The answer, says the author of this patent, is to find the impulse response of the room at the remote and subtract the ambient sounds from the input. Next, the user's voice is deconvolved with the room impulse response and then voice recognition can be attempted. No concrete results are given. Again, the examiner has researched the prior art.—MK

6,601,029

43.72.Ne VOICE PROCESSING APPARATUS

John Brian Pickering, assignor to International Business Machines Corporation
29 July 2003 (Class 704/257); filed in the United Kingdom 11 December 1999

An algorithm finds uncertainties in the outputs of a speech recognizer



and notifies the user of the uncertain parts.—HHN

6,604,073

43.72.Ne VOICE RECOGNITION APPARATUS

Shoutarou Yoda, assignor to Pioneer Corporation
5 August 2003 (Class 704/231); filed in Japan 12 September 2000

A method for speech recognition in noisy environments captures an image from the user's mouth to aid in determining the start and end points of the spoken phrase.—HHN

6,606,598

43.72.Ne STATISTICAL COMPUTING AND REPORTING FOR INTERACTIVE SPEECH APPLICATIONS

Mark A. Holthouse *et al.*, assignors to Speech Works International, Incorporated
12 August 2003 (Class 704/275); filed 21 September 1999

An algorithm reports statistical information that describes the performance of an interactive speech application.—HHN

6,608,249

43.72.Ne AUTOMATIC SOUNDTRACK GENERATOR

Alain Georges, assignor to dBTech SARL
19 August 2003 (Class 84/609); filed 21 May 2002

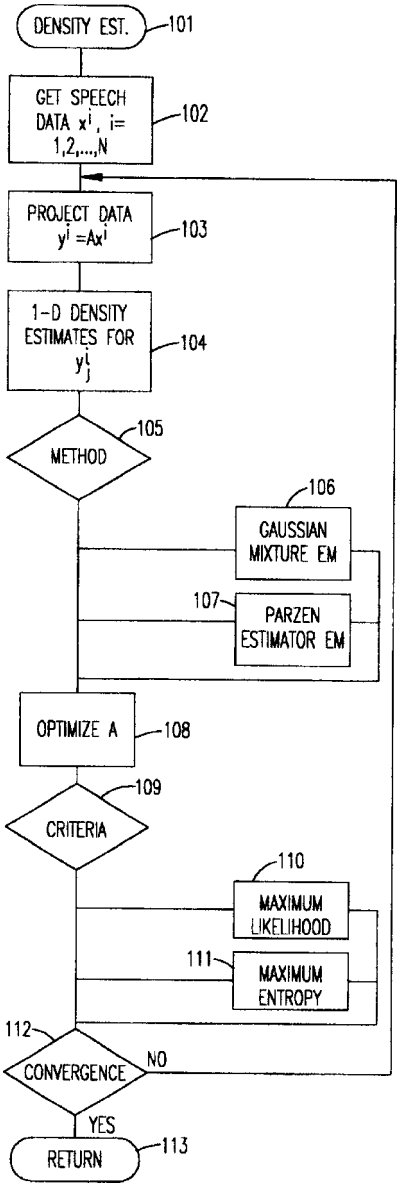
This patent describes a device for sound tracking by integrating a sound generator module in a video recorder/player for sound mixing and dubbing purposes.—HHN

6,609,094

43.72.Ne MAXIMUM ENTROPY AND MAXIMUM LIKELIHOOD CRITERIA FOR FEATURE SELECTION FROM MULTIVARIATE DATA

Sankar Basu *et al.*, assignors to International Business Machines Corporation
19 August 2003 (Class 704/240); filed 22 May 2000

This patent describes a method for speech recognition by projecting the high-dimensional data into lower dimensional subspaces, estimating the univariate power density function (PDF), maximizing the entropy or likelihood, and then reconstructing the PDF in the original high dimension.—HHN

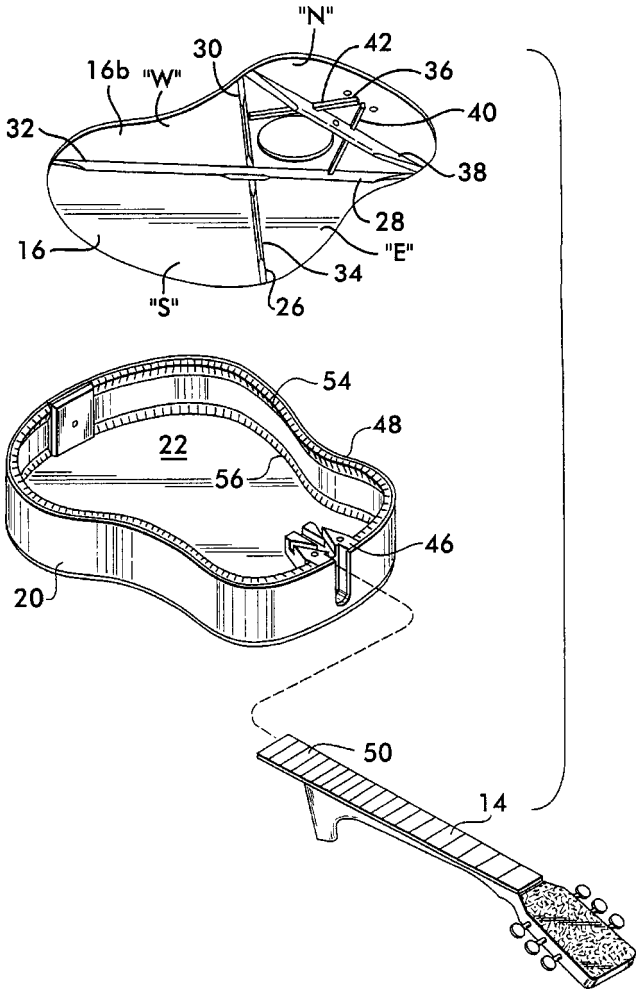


6,605,766

43.75.Gh ACOUSTIC GUITAR ASSEMBLY

Timothy A. Teel, assignor to C. F. Martin & Company, Incorporated
12 August 2003 (Class 84/291); filed 22 May 2001

To provide a ''high quality, durable acoustic guitar which is at least partly constructed from alternative nonwooden materials which provide both



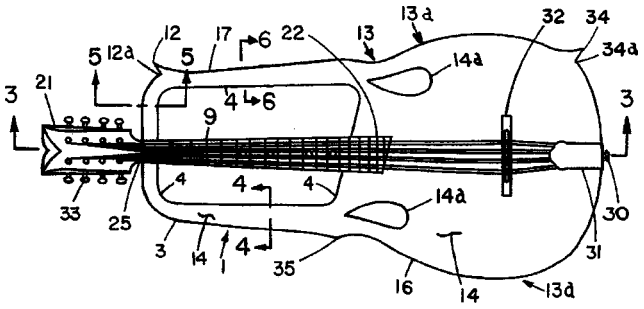
a novel decorative appearance and superior acoustic performance,'' Martin Guitar proposes using a thin metal soundboard **16** with laminate sidewalls **20** and backboard **22**.—MK

6,608,247

43.75.Gh STRINGED MUSICAL INSTRUMENT WITH SOUNDBOX EXTENSION

Harvard Jasper Bryan, Magnolia, Arkansas
19 August 2003 (Class 84/291); filed 30 October 2001

According to the patent author, a mandolin suffers from a number of deficiencies including the ''power'' of the sound and less flexion resulting in



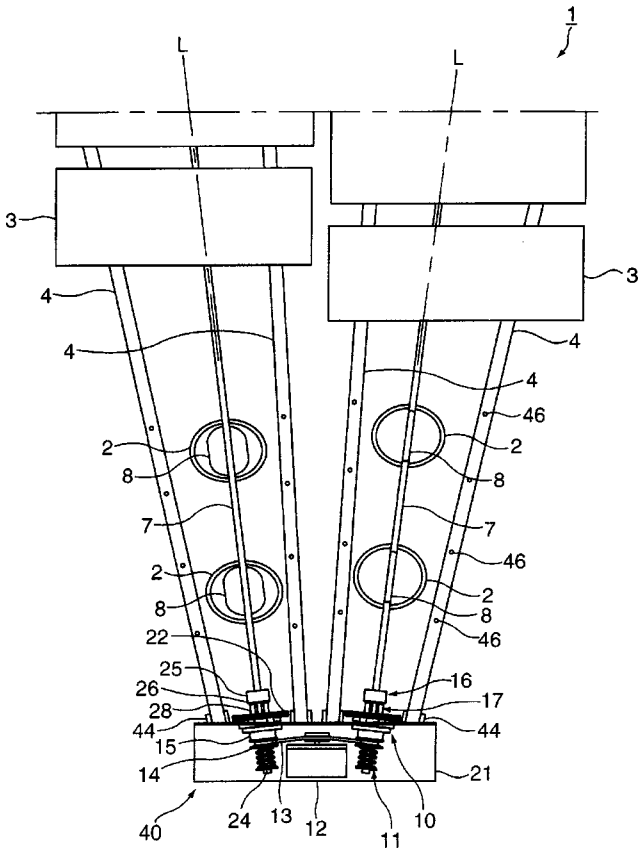
tuning instabilities. The proposed cure is an extended sound box that resembles a lyre. Issues of body resonance are not addressed.—MK

6,596,932

43.75.Kk VIBRAPHONE HAVING IMPROVED DRIVE OF FANS

André Ansfried J. Adams, assignor to Adams Paukenfabriek B.V.
22 July 2003 (Class 84/402); filed in the Netherlands 16 November 2000

The wonderful shimmering sound of the vibraphone is due to the set of butterfly valves **8** located between bars **3** and resonance tubes **2**. The author



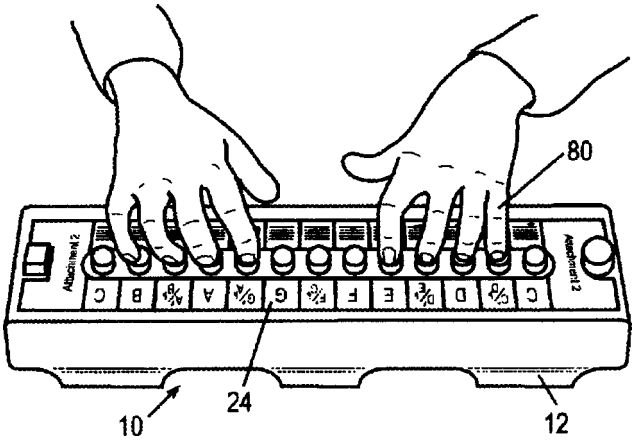
proposes a slip coupling to prevent the musician from damaging valuable body parts as well as the belt drive **14** (driven by motor **12**).—MK

6,605,767

43.75.St MUSIC LEARNING SYSTEM INCLUDING INSTRUMENT AND METHOD

Ethan Fiks and Peter Sippach, assignors to Music Path Company, Incorporated
12 August 2003 (Class 84/423 R); filed 6 November 2001

Teachingbeginningmusicusingtherecorderismademorecomplicated by the difficulty of producing even breath control, not to mention the hygenic issues of sharing a single instrument. Accordingly, the inventors be-



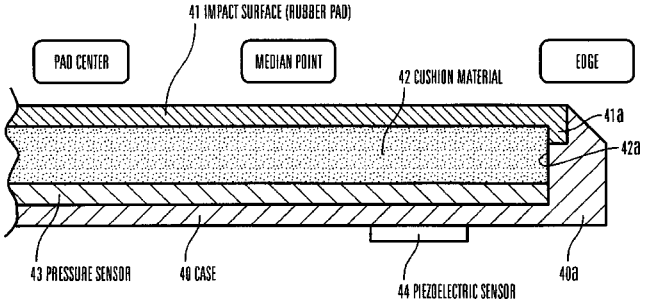
lieve a keyboard instrument with switchable labels is preferable. Issues of woodwind fingering are eliminated as well. But then why not use a single octave piano keyboard? At least the skill is transferable to another instrument.—MK

6,601,436

43.75.Wx APPARATUS AND METHOD FOR DETECTING AND PROCESSING IMPACTS TO AN ELECTRONIC PERCUSSION INSTRUMENT

Kenji Senda and Yasunobu Miyamoto, assignors to Roland Corporation
5 August 2003 (Class 73/12.09); filed 26 January 2001

Roland Corporation has managed to extract 45 claims after describing an electronic drum pad. As shown, they propose two detectors: one pressure sensor to detect hits and a piezoelectric sensor to help detect location. It's



very strange that no reference was made to Max Mathews' electronic drum. After all, it was published in the *Computer Music Journal* only 23 years ago.—MK

6,605,768

43.75.Wx MUSIC-SIGNAL COMPRESSING/ DECOMPRESSING APPARATUS

Toshihiko Date *et al.*, assignors to Matsushita Electric Industrial Company, Limited
12 August 2003 (Class 84/604); filed in Japan 6 December 2000

In this conceptual patent, a sound source is first separated into different streams of different importance (of course, the algorithm is not disclosed, described, or discussed). Next, depending on the importance, the streams can be coded in different ways (e.g., MIDI for music, CELP for voice, etc.) and then transmitted or stored.—MK

6,605,769

43.75.Wx MUSICAL INSTRUMENT DIGITAL RECORDING DEVICE WITH COMMUNICATIONS INTERFACE

Henry E. Juszkiewicz, assignor to Gibson Guitar Corporation
12 August 2003 (Class 84/609); filed 5 October 2001

Incredibly, the Gibson Guitar company believes that a DSP microprocessor and a USB port are not only novel but also worth patenting. Once again, the patent examiner supplied all the prior art citations.—MK

6,607,488

43.80.Qf MEDICAL DIAGNOSTIC ULTRASOUND SYSTEM AND METHOD FOR SCANNING PLANE ORIENTATION

John I. Jackson and John A. Hossack, assignors to Acuson Corporation
19 August 2003 (Class 600/443); filed 2 March 2000

This diagnostic ultrasound system indicates the position and orientation of a scan plane relative to a patient. The image as presented on a display follows the orientation of a transducer. The orientation may provide display of an image where the transducer is not in an up or down position, but is rotated away from the vertical of the display. As an alternative, a two- or three-dimensional graphic generic representation is provided with a scan plane represented as a polygon or image rendering within a generic representation to display relative positioning of the scan plane with respect to the patient.—DRR

6,607,502

43.80.Sh APPARATUS AND METHOD INCORPORATING AN ULTRASOUND TRANSDUCER ONTO A DELIVERY MEMBER

Mark A. Maguire and James C. Peacock III, assignors to Atrionix, Incorporated
19 August 2003 (Class 604/22); filed 5 November 1999

This patent involves a device that consists of an ultrasound transducer assembly mounted on a catheter shaft in order to ultrasonically couple to a region of tissue within the body of a patient and, more specifically, to couple to a circumferential region of tissue at a location where a pulmonary vein extends from the atrium. Mounting flanges extend from either end of the transducer and are mounted at two locations along the catheter shaft so that the transducer is suspended over the catheter shaft between those mounting locations. In an overall tissue coupling system, a cylindrical ultrasound transducer is mounted onto a catheter shaft by the use of such mounting flanges in such a manner that the ultrasound transducer is enclosed within a balloon that is also provided in the catheter.—DRR

6,607,487

43.80.Vj ULTRASOUND IMAGE GUIDED ACETABULAR IMPLANT ORIENTATION DURING TOTAL HIP REPLACEMENT

John Chang *et al.*, assignors to The Regents of the University of California
19 August 2003 (Class 600/437); filed 23 January 2001

An ultrasound probe is used to survey the pelvic bone around the acetabulum. An adapter cup is then positioned in the acetabular socket between the ultrasonic imaging head and the socket. Marks are made in the socket relative to the pelvis. The marks are used to locate the ultrasonic adapter cup and acetabular implant in the joint socket.—RCW

6,607,489

43.80.Vj FOCUS CORRECTION FOR ULTRASOUND IMAGING THROUGH MAMMOGRAPHY COMPRESSION PLATE

Ralph Thomas Hoctor and Kai Erik Thomenius, assignors to General Electric Company
19 August 2003 (Class 600/443); filed 5 April 2001

Ray tracing is used to adjust ultrasonic imaging instrument transmit and receive time delays based on compression plate thickness and sound speed in the plate to mitigate spherical aberration due to the presence of the compression plate.—RCW

Comment on "Background noise in piezoresistive, electret condenser, and ceramic microphones" [J. Acoust. Soc. Am. 113, 3179–3187 (2003)] (L)

Richard Brander[a)]
RB Research, 1429 S. 59th Avenue, Cicero, Illinois 60804

(Received 8 August 2003; revised 31 October 2003; accepted 5 November 2003)

This letter comments on the paper ''Background noise in piezoresistive, electret condenser, and ceramic microphones'' [J. Acoust. Soc. Am. **113**, 3179–3187 (2003)], which found low-frequency-weighted noise in microphones. One possible mechanism for this type of noise is proposed. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639329]

PACS numbers: 43.38.Ar, 43.38.Kb [AJZ] Pages: 465–466

In a recent paper,[1] Zuckerwar, Kuhn, and Serbyn found low-frequency-weighted noise generated by diaphragm motion in piezoresistive and electret condenser microphones. This finding extended previous work[2] in which a similar type of noise was found in condenser microphones of various sizes. However, no physical mechanism for the generation of this type of noise was identified. The purpose of this communication is to propose one possible mechanism for this type of noise.

A generalized model for a pressure microphone is shown in Fig. 1. In this figure, C_A represents the acoustic compliance of the diaphragm and moving system, M_A represents the acoustic mass of the diaphragm and moving system, and R_A represents the acoustic damping resistance. P represents the input acoustic pressure and V_0 represents the open circuit output voltage. These elements correspond to similar elements in Fig. 3 and Fig. 6 of Ref. 1. Additional elements in Fig. 1 are C_B, the acoustic compliance of the back cavity volume, and R_B, the acoustic resistance of the pressure relief vent for the back cavity. A closed back cavity is needed to shield the back side of the diaphragm from the input acoustic pressure, and a pressure relief vent is needed to accommodate changes in barometric pressure without static offset of the diaphragm. These elements have been incorporated in previous microphone models.[3]

Two noise sources are identified in Fig. 1. These are *nA*, the Johnson noise associated with R_A, and *nB*, the Johnson noise associated with R_B. The power spectral densities of these noise sources are $S_{nA}=4kTR_A$ and $S_{nB}=4kTR_B$, respectively. Z_1 is the acoustic impedance looking into the network to the right of the dotted line and T_1 is the transfer function V_0/P_1 of this network.

Let V_{0P}, V_{0A}, and V_{0B} represent the output voltage produced by the input pressure P, the noise source *nA*, and the noise source *nB*, respectively.

$$V_{0P}=\frac{sC_BR_B}{sC_BR_B+1+R_B/Z_1}\cdot P\equiv T_P\cdot P, \tag{1}$$

$$V_{0A}=\frac{sC_BR_B+1}{sC_BR_B+1+R_B/Z_1}\cdot nA, \tag{2}$$

$$V_{0B}=\frac{1}{sC_BR_B+1+R_B/Z_1}\cdot nB. \tag{3}$$

The equivalent input noise pressure due to noise sources *nA* and *nB* is given by

$$P_{inA}=\frac{V_{0A}}{T_P}=\left(1+\frac{1}{sC_BR_B}\right)\cdot nA, \tag{4}$$

$$P_{inB}=\frac{V_{0B}}{T_P}=\left(\frac{1}{sC_BR_B}\right)\cdot nB. \tag{5}$$

The equivalent input noise power spectral density is given by

$$S_{inA}=\left(1+\frac{1}{4\pi^2f^2C_B^2R_B^2}\right)\cdot 4kTR_A, \tag{6}$$

$$S_{inB}=\left(\frac{1}{4\pi^2f^2C_B^2R_B^2}\right)\cdot 4kTR_B. \tag{7}$$

As an example, consider a microphone with the following parameters: $C_A=10^{-14}$ m^5/N, resonant frequency $f_R=20$ kHz, quality factor $Q=1$, $C_B/C_A=5$, and pressure relief corner frequency $f_L=2$ Hz.

$$R_A=\frac{1}{2\pi\cdot f_RC_AQ}=7.95\times10^8\ \text{Ns/m}^5, \tag{8}$$

$$R_B=\frac{1}{2\pi\cdot f_LC_B}=1.59\times10^{12}\ \text{Ns/m}^5. \tag{9}$$

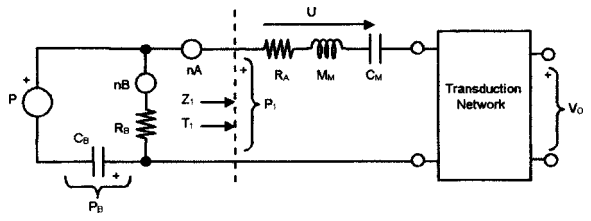


FIG. 1. Generalized model for a pressure microphone.

[a)]Electronic mail: rdbrander@cs.com

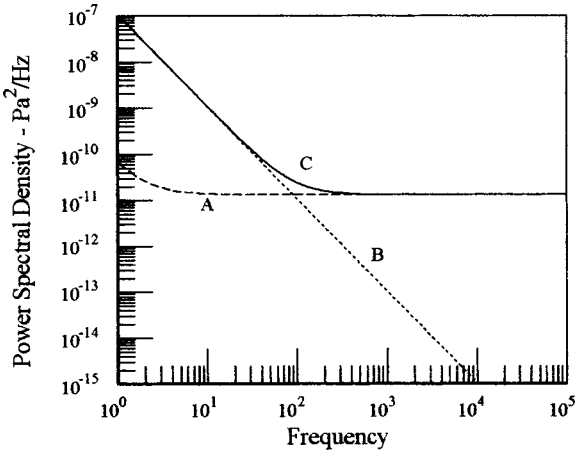


FIG. 2. Equivalent input noise power spectral density. Curve A: due to damping resistance R_A. Curve B: due to vent resistance R_B. Curve C: sum of A and B.

Figure 2 shows the equivalent input noise power spectral density for the noise due to R_A, the noise due to R_B, and the sum of these. It can be seen that a vent to a back cavity is one possible explanation for acoustically generated low-frequency noise.

[1] A. J. Zuckerwar, T. R. Kuhn, and R. M. Serbyn, ''Background noise in piezoresistive, electret condenser, and ceramic microphones,'' J. Acoust. Soc. Am. **113**, 3179–3187 (2003).

[2] A. J. Zuckerwar and K. C. T. Ngo, ''Measured 1/f noise in the membrane motion of condenser microphones,'' J. Acoust. Soc. Am. **95**, 1419–1425 (1994).

[3] E. Frederiksen, ''Microphone system for extremely low sound levels,'' Bruel & Kjaer Technical Review 1984 n3 15–22 (1984).

Acoustical wave propagator for time-domain flexural waves in thin plates

S. Z. Peng and J. Pan[a)]
Centre for Acoustics, Dynamics and Vibration, School of Mechanical Engineering, The University of Western Australia, Crawley, WA 6009, Australia

(Received 12 September 2002; revised 24 September 2003; accepted 8 November 2003)

In this paper, an explicit acoustical wave propagator technique is introduced to describe the time-domain evolution of acoustical waves in two-dimensional plates. A combined scheme with Chebyshev polynomial expansion and fast Fourier transformation is used to implement the operation of the acoustical wave propagator. Through this operation, the initial wave packet at $t=0$ is mapped into the wave packet at any instant $t>0$. By comparison of the results of the exact analytical solution and the Euler numerical method, we find that this new Chebyshev–Fourier scheme is highly accurate and computationally effective in predicting the acoustical wave propagation in thin plates. This method offers an opportunity for future study of dynamic stress concentration and time-domain energy flow in coupled structures. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639905]

PACS numbers: 43.20.Bi, 43.20.Fn, 43.40.Dx [ANN] Pages: 467–474

I. INTRODUCTION

Direct modeling of time-domain wave propagation provides an important tool in studying transient wave behavior at discontinuities. Intimate details of the wave properties can be revealed by examining the time-domain propagation of the wave packet. Much effort has been made in developing the finite difference and finite element models for time-domain analysis of acoustic waves.[1–8] There is still an urgent need for effective and accurate methods to investigate a range of interesting acoustical problems, including dynamic stress concentration at structural discontinuities and transient energy exchange across the coupling boundaries. Tal-Ezer and Kosloff[9] pioneered a new propagation scheme for solving the time dependent Schrödinger equation in molecular dynamics. This scheme was regarded as a global propagator method, especially applicable to an explicit time-independent Hamiltonian operator. It allows very long time steps, and often a single time step can complete the calculation. This scheme has been successfully employed in the field of quantum chemistry and atomic physics. Considerable work has been done over the last two decades for the resolution of elastic wave propagation phenomena. Relevant to the theory and application of wave propagators, we find two main streams of research. The work done by Olsson,[10] Karlsson,[11] and Billger and Folkow[12] on wave propagators (spatial propagators) focuses on the properties of the propagators defined on the basis of the wave splitting technique, and the relations between the method of propagators and the invariant imbedding method and the Green function approach. On the other hand, the acoustical wave propagator (temporal propagator),[13] defined on the basis of Euler's field equations, leads directly to the fast and accurate solution of the evolution of the initial disturbance by employing fast converging expansion functions.

The technique of the acoustical wave propagator (AWP) was developed by comparing the difference in operators for Schrödinger equations and sound wave equations.[13] To implement the operation of the AWP on an initial state of wave motion, the AWP was expanded by the Chebyshev polynomial in the time domain. Meanwhile, the spatial gradient in each polynomial term was evaluated by the Fourier transformation scheme. Analysis and numerical examples demonstrated that this Chebyshev–Fourier scheme was highly accurate and computationally effective in predicting the time-domain acoustical wave propagation and scattering. However, in Pan and Wang's paper, AWP is only applied to one-dimensional sound and structural wave packet evolution. The aim of this work is to extend the acoustical wave propagator approach to investigate the evolution of acoustic wave packets in two-dimensional thin plates.

II. THEORY OF THE ACOUSTICAL WAVE PROPAGATOR

Plates are important components of many practical engineering structures. The governing equations of a thin plate are based on Kirchhoff's hypothesis.[14] They are relevant for small deflections of a flat plate. Very often, the thickness of the plate is much smaller than sizes in other dimensions, so waves in thin plates are usually treated as two-dimensional waves.[15] A plate may be subjected to a distributed external pressure load $P(x,y,t)$ as a function of time. The moments and shearing forces due to internal stress resultants in the plate are generated to balance the external load together with inertia forces. Two kinds of moments are involved: bending moments, M_x and M_y, and twisting moments, M_{xy} and M_{yx}.

The resultant bending and twisting moments are related to the plate velocity V by

a)Electronic mail: pan@mech.uwa.edu.au

$$\frac{\partial M_x}{\partial t}=-D\left[\frac{\partial^2 V}{\partial x^2}+\upsilon\frac{\partial^2 V}{\partial y^2}\right],\quad \frac{\partial M_y}{\partial t}=-D\left[\frac{\partial^2 V}{\partial y^2}+\upsilon\frac{\partial^2 V}{\partial x^2}\right],$$

$$\frac{\partial M_{xy}}{\partial t}=-D(1-\upsilon)\left[\frac{\partial^2 V}{\partial x\partial y}\right],\quad \frac{\partial M_{yx}}{\partial t}=D(1-\upsilon)\left[\frac{\partial^2 V}{\partial x\partial y}\right], \quad (1)$$

where $D=2Eh^3/3(1-\upsilon^2)$ is the flexural rigidity of the plate, and $2h$, υ, and E are the thickness, Poisson's ratio, and Young's modulus of the plate, respectively.

Considering the force equilibrium in the z direction and moment equilibrium about the x axis and y axis, the shearing forces Q_x and Q_y are related to the moments as

$$Q_x=\frac{\partial M_x}{\partial x}+\frac{\partial M_{xy}}{\partial y},\quad Q_y=\frac{\partial M_y}{\partial y}+\frac{\partial M_{xy}}{\partial x}. \quad (2)$$

Thus the force equilibrium condition in the z direction and Eq. (2) leads to the fourth-order Euler–Bernoulli flexural wave equation for a 2D thin plate,

$$\frac{\partial V}{\partial t}=\frac{1}{2\rho_s h}\left(\frac{\partial Q_x}{\partial x}+\frac{\partial Q_y}{\partial y}\right)+\frac{P(x,y,t)}{2\rho_s h}$$
$$=\frac{1}{2\rho_s h}\left(\frac{\partial^2 M_x}{\partial x^2}+2\frac{\partial^2 M_{xy}}{\partial x\partial y}+\frac{\partial^2 M_y}{\partial y^2}\right)+\frac{P(x,y,t)}{2\rho_s h}, \quad (3)$$

where ρ_s is the mass per unit area of the plate. It is noted that the external load $P(x,y,t)$ may include static loading $2h\rho_s g$ due to gravity force and dynamic load. The static loading only generates a static response. In this paper, only the time-dependent response to initial conditions is discussed. For this case, $P(x,y,t)=0$.

To derive the acoustic wave propagator for flexural waves in a thin plate, we select a state vector ϕ consisting of velocity V, bending moments $M_x(x,y,t)$, $M_y(x,y,t)$ and twisting moment $M_{xy}(x,y,t)$ of the plate element. Combining Eqs. (3) and (1), we have the system state equation

$$\frac{\partial}{\partial t}\begin{bmatrix} V \\ M_x \\ M_y \\ M_{xy}\end{bmatrix}=-\hat{\mathbf{H}}\begin{bmatrix} V \\ M_x \\ M_y \\ M_{xy}\end{bmatrix}, \quad (4)$$

where

$$\hat{\mathbf{H}}=\begin{bmatrix} 0 & -\frac{1}{2\rho_s h}\frac{\partial^2}{\partial x^2} & -\frac{1}{2\rho_s h}\frac{\partial^2}{\partial y^2} & -\frac{1}{\rho_s h}\frac{\partial^2}{\partial x\partial y} \\ D\left(\frac{\partial^2}{\partial x^2}+\upsilon\frac{\partial^2}{\partial y^2}\right) & 0 & 0 & 0 \\ D\left(\frac{\partial^2}{\partial y^2}+\upsilon\frac{\partial^2}{\partial x^2}\right) & 0 & 0 & 0 \\ D(1-\upsilon)\frac{\partial^2}{\partial x\partial y} & 0 & 0 & 0\end{bmatrix}. \quad (5)$$

Integrating Eq. (4) with respect to time, we then obtain

$$\phi(x,y,t)=e^{-(t-t_0)\hat{\mathbf{H}}}\phi(x,y,t_0), \quad (6)$$

where $e^{-(t-t_0)\hat{\mathbf{H}}}$ is defined as the acoustical wave propagator for flexural waves in thin plates. In general, wave propagators are operators that map one field at one position and time to another.[11,16] However, their mathematical forms vary with applications. Propagators defined directly from the wave equation are often asymmetrical.[9,13] While these are defined after certain transformations, such as the wave splitting transformation, the resultant matrix propagator could be symmetrical if the media for wave propagation is isotropic.[6,11] It is noted that the operator $e^{-(t-t_0)\hat{\mathbf{H}}}$ is a temporal propagator, not a spatial propagator, as it applies to a fixed point in space and advances in time. The state vector $\phi(x,y,t)$ of an acoustical wave at any time t can be evaluated by the operation of the acoustical wave propagator (AWP) $e^{-(t-t_0)\hat{\mathbf{H}}}$ acting upon the initial state vector $\phi(x,y,t_0)$. The boundary and spatial variation of the acoustical media are readily included in the system operator $\hat{\mathbf{H}}$ by defining the parameters in $\hat{\mathbf{H}}(D,\upsilon,\rho_s,h)$ as a function of position.

III. IMPLEMENTATION OF AWP

A. Chebyshev polynomial expansion of AWP

From Eq. (6) and Eq. (2), we are able to obtain complete information about the whole system in the time-domain ($W(x,y,t)$, $V(x,y,t)$, M_x, M_y, M_{xy}, Q_x, and Q_y). The exponential propagator needs to be expanded so that its operation on the initial wave packet can be evaluated. Different expansions to the propagator along with the technique used to calculate the spatial derivatives lead to different time-evolution schemes. The simplest scheme is based on the Taylor expansion, known as Euler's method. Recent work on the time-domain Schrödinger equation indicates that the Chebyshev polynomial expansion (CPE) scheme has some distinct

features. First, it allows the use of very long time steps: in some cases, a single time step completes the whole calculation. When the wave propagator for the time-domain Schrödinger equation is expanded by the Chebyshev polynomials, there is virtually no limit to the length of the time step used in the expansion. Modern computing facilities allow the use of many expansion terms to support a large step for the wave propagation. As a result, a single event of scattering can be calculated using a single time step. Second, the expansion coefficients of Chebyshev polynomials decay exponentially when the order of the coefficient function is sufficiently larger than its argument. The acoustical wave propagator is in a matrix form. Real Chebyshev polynomials are used in the expansion of the acoustical wave propagator, and matrix operator algebra is required for its implementation. As mentioned above, $\hat{\mathbf{H}}$ is an operator with a valued matrix so that λ is an eigenvalue operator of the system operator. The range of an operator is defined by the maximum and minimum values (or eigenvalues for matrix operator) of the operation result upon a function. If a variable is divided by its upper limit (maximum value), the resultant normalized variable will be restricted to $[-1,1]$. The real Chebyshev polynomials are defined in the ranges of $[-1,1]$. To ensure a convergent Chebyshev polynomial expansion, the system operation $\hat{\mathbf{H}}$ needs to be normalized by

$$\hat{\mathbf{H}}' = \frac{\hat{\mathbf{H}}}{\lambda_{\max}}, \tag{7}$$

where $\lambda_{\max}$ is the maximum eigenvalue λ of the system operator.

As described by Ref. 13, the acoustical wave propagator can be expanded in Chebyshev polynomials of the first kind as

$$\mathbf{e}^{-(t-t_0)\hat{\mathbf{H}}} = \mathbf{e}^{-R\hat{\mathbf{H}}'} = \sum_{n=0}^{\infty} a_n(R) T_n(\hat{\mathbf{H}}'), \tag{8}$$

where a_0, $a_1,\ldots,a_{n-1}$, a_n, $a_{n+1},\ldots$ are the expansion coefficients and T_n is called the Chebyshev polynomial of the order n; $R=(t-t_0)\lambda_{\max}$; $a_n(R)=2I_n(R)$ except $a_0(R)=I_0(R)$ and $I_n(R)$ is the nth-order modified Bessel function of the first kind. The zero- and first-order Chebyshev polynomials are defined as $T_0(\hat{\mathbf{H}}')=\mathbf{I}$ and $T_1(\hat{\mathbf{H}}')=\hat{\mathbf{H}}'$, and the rest can be calculated by the following recursive relation:

$$T_{n+1}(\hat{\mathbf{H}}') = 2\hat{\mathbf{H}}' T_n(\hat{\mathbf{H}}') - T_{n-1}(\hat{\mathbf{H}}'). \tag{9}$$

Equation (6) can be rewritten as

$$\begin{aligned}\phi(x,y,t) &= e^{-(t-t_0)\hat{\mathbf{H}}}\phi(x,y,t_0) \\ &= \sum_{n=0}^{\infty} a_n(R) T_n(\hat{\mathbf{H}}')\phi(x,y,t_0).\end{aligned} \tag{10}$$

The AWP operation therefore becomes the calculation of the operation of the normalized system operator $\hat{\mathbf{H}}'$ on the initial state vector, which mainly involves the evaluation of spatial derivatives.

Utilizing the Taylor expansion of $e^{-\hat{\mathbf{H}}\,dt}$, the acoustical wave propagator is given in

$$\begin{aligned}\phi(x,y,t+dt) &= e^{-\hat{\mathbf{H}}\,dt}\phi(x,y,t) \\ &= (1-\hat{\mathbf{H}}\,dt+\cdots)\phi(x,y,t).\end{aligned} \tag{11}$$

If only the first two terms in the Taylor expansion of the AWP are considered, then an explicit second-order propagation scheme is obtained:

$$\phi(x,y,t+dt) = (1-\hat{\mathbf{H}}\,dt)\phi(x,y,t). \tag{12}$$

B. Fourier transformation scheme for spatial derivatives

The complex Fourier transformation is used to evaluate spatial derivatives of $\phi(x,y,t)$.

$$\begin{aligned}\mathcal{F}\left[\frac{\partial^N}{\partial x^N}\phi(x,y,t)\right] &= (jk_x)^N \mathcal{F}[\phi(x,y,t)], \\ \mathcal{F}\left[\frac{\partial^N}{\partial y^N}\phi(x,y,t)\right] &= (jk_y)^N \mathcal{F}[\phi(x,y,t)],\end{aligned} \tag{13}$$

where $\mathcal{F}[\ ]$ represents the Fourier transformation and k_x and k_y are the wave numbers as in $e^{jk_x x}$ and $e^{jk_y y}$, respectively. Therefore, the calculation of spatial derivatives of function $\phi(x,y,t)$ is obtained by the following inverse Fourier transformations:

$$\begin{aligned}\frac{\partial^N}{\partial x^N}\phi(x,y,t) &= \mathcal{F}^{-1}\{(jk_x)^N \mathcal{F}[\phi(x,y,t)]\}, \\ \frac{\partial^N}{\partial y^N}\phi(x,y,t) &= \mathcal{F}^{-1}\{(jk_y)^N \mathcal{F}[\phi(x,y,t)]\}.\end{aligned} \tag{14}$$

The accuracy of the Fourier transformation method for evaluating spatial derivatives has been examined.[13] When the spatial sampling intervals of Δx and Δy are given, the discrete Fourier expansion of a wave packet supports a range of wave-number components with a limit specified by the maximum wave number:

$$k_{B_{\max}} = \sqrt{\left(\frac{\pi}{\Delta x}\right)^2 + \left(\frac{\pi}{\Delta y}\right)^2}, \tag{15}$$

where $k_{B_{\max}}$ represents the maximum bending wave number. For flexural waves in a thin plate, the normalization factor R is

$$\begin{aligned}R &= (t-t_0)\lambda_{\max} \\ &= \sqrt{\frac{Eh^2}{3\rho_s(1-\nu^2)}}\left\{\left(\frac{\pi}{\Delta x}\right)^2 + \left(\frac{\pi}{\Delta y}\right)^2\right\}(t-t_0).\end{aligned} \tag{16}$$

Components with wave number above the limit specified by Eq. (15) will not be included in the numerical model. The convergence and accuracy properties of the spatial derivative calculation using Fourier transformation (wave number transformation) method have been discussed analytically and numerically by Pan and Wang[13] and Wang.[17]

C. An exact solution

To demonstrate the application of the AWP, first of all, the exact analytical solution of the initial wave packet in a thin plate is presented below. Assume the following initial displacement and velocity:

$$W(r,0)=f(r), \quad \frac{\partial W}{\partial t}(r,0)=0. \tag{17}$$

For a specific initial waveform, $W(r,0)=f_0 e^{-(r^2/4\sigma^2)}$, where f_0 is a constant, σ is a Gaussian factor. The displacement has the following analytical solution:[18]

$$W(r,t)=\frac{f_0}{1+\tau^2}e^{[-\mu^2/(1+\tau^2)]}\left\{\cos\left(\frac{\mu^2\tau}{1+\tau^2}\right)+\tau\sin\left(\frac{\mu^2\tau}{1+\tau^2}\right)\right\}, \tag{18}$$

where $\tau=(Eh^2/3(1-\nu^2)\rho_s)^{1/2}(t/\sigma^2)$; $\mu=r/2\sigma$.

The first-order derivative of $W(r,t)$ is given by

$$\frac{\partial W(r,t)}{\partial r}=\frac{f_0\mu}{\sigma(1+\tau^2)^2}e^{[-\mu^2/(1+\tau^2)]}\times\left\{(\tau^2-1)\cos\left(\frac{\mu^2\tau}{1+\tau^2}\right)-2\tau\sin\left(\frac{\mu^2\tau}{1+\tau^2}\right)\right\}. \tag{19}$$

The second-order derivative of $W(r,t)$ is given by

$$\frac{\partial^2 W(r,t)}{\partial r^2}=\frac{f_0}{\sigma^2(1+\tau^2)^3}e^{[-\mu^2/(1+\tau^2)]}\times\left\{\left(\frac{\tau^4-6\mu^2\tau^2+2\mu^2-1}{2}\right)\cos\left(\frac{\mu^2\tau}{1+\tau^2}\right)+\tau(3\mu^2-\mu^2\tau^2-1-\tau^2)\sin\left(\frac{\mu^2\tau}{1+\tau^2}\right)\right\}. \tag{20}$$

The first-order and second-order derivatives of $W(r,t)$ lead to the exact expressions of the bending and twisting moments in polar coordinates,

$$M_r=-D\left[\frac{\partial^2 W}{\partial r^2}+\upsilon\frac{1}{r}\frac{\partial W}{\partial r}\right]=-\frac{Df_0}{2\sigma(1+\tau^2)^2}e^{[-\mu^2/(1+\tau^2)]}\times\left\{A_1\cos\left(\frac{\mu^2\tau}{1+\tau^2}\right)+A_2\sin\left(\frac{\mu^2\tau}{1+\tau^2}\right)\right\}, \tag{21}$$

where

$$A_1=\frac{r(\tau^4-6\tau^2\mu^2+2\mu^2-1)+2\upsilon\mu\sigma(\tau^4-1)}{\sigma(1+\tau^2)r},$$

$$A_2=\frac{2\tau r(3\mu^2-\mu^2\tau^2-1-\tau^2)-4\upsilon\mu\sigma\tau(1+\tau^2)}{\sigma(1+\tau^2)r}$$

and

$$M_\theta=-D\left[\upsilon\frac{\partial^2 W}{\partial r^2}+\frac{1}{r}\frac{\partial W}{\partial r}\right]=-\frac{Df_0}{2\sigma(1+\tau^2)^2}e^{[-\mu^2/(1+\tau^2)]}\times\left\{A_3\cos\left(\frac{\mu^2\tau}{1+\tau^2}\right)+A_4\sin\left(\frac{\mu^2\tau}{1+\tau^2}\right)\right\}, \tag{22}$$

where

$$A_3=\frac{\upsilon r(\tau^4-6\tau^2\mu^2+2\mu^2-1)+2\mu\sigma(\tau^4-1)}{\sigma(1+\tau^2)r},$$

$$A_4=\frac{2\upsilon\tau r(3\mu^2-\mu^2\tau^2-1-\tau^2)-4\mu\sigma\tau(1+\tau^2)}{\sigma(1+\tau^2)r}.$$

This allows an assessment of the prediction accuracy of the Chebyshev–Fourier scheme developed in the preceding section and of the Euler method described above.

D. Convergence property of Chebyshev–Fourier scheme

The convergence property of Chebyshev–Fourier scheme is determined by the convergence property of the Chebyshev expansion of Eq. (10). If the expansion coefficients A_k in the infinite series $s=\Sigma_{k=0}^{\infty}A_kT_k$ approach zero when $k>n$, and the truncation error $|\Sigma_{k=0}^{\infty}A_kT_k-\Sigma_{k=0}^{n}A_kT_k|$ is less than ϵ, where ϵ is a small positive value dependent upon n, then the finite series $\Sigma_{k=0}^{n}A_kT_k$ converges to s with a maximum truncation error of ϵ.

The Chebyshev expansion in Eq. (10) has a property of $A_k\to 0$ when $k\to\infty$.[19] Theoretically we can obtain any degree of accuracy by choosing a suitable n. Practically, the accuracy of the expansion [Eq. (10)] is also dependent on the round-off error in the calculation of the expansion coefficients. It becomes clear that we have to examine the effect of the round-off error on the convergence property of the expansion. The round-off error is due to the finite precision of the computer. For example, a double precision will have a ''dynamic range'' of 10^{16}. If the largest value of the coefficient is of the order 10^{16}, then in the calculation any significant value of all the coefficients less than the order of 1 will be truncated (or rounded-off).

Figure 1 shows the numerical values of $I_n(R)$ as functions of n and R. For a given ''time step'' R, the coefficients of the expansion decay almost exponentially with the increase of n. This result supports the convergence requirement of the Chebyshev expansion as in Eq. (10). Furthermore, Fig. 1 shows the large variation among the numerical values of the coefficients, which indicates the potential round-off error problem in the accurate calculation of the expansion. To investigate the effect of the round-off error on prediction accuracy, we first define a number $n_{\min}$ so that the terms with n

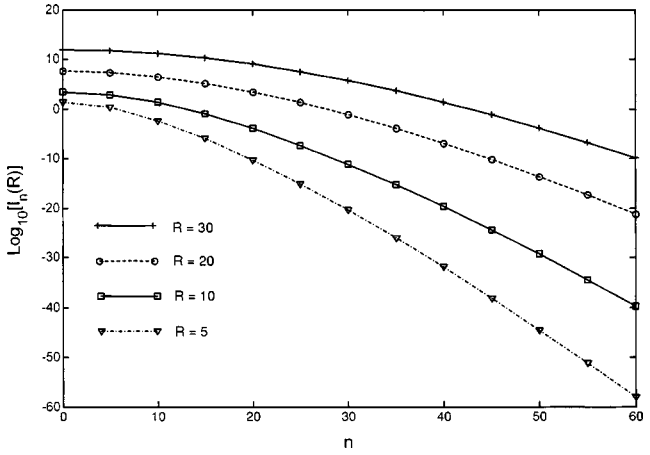


FIG. 1. The convergence properties of $I_n(R)$ as functions of n and R.

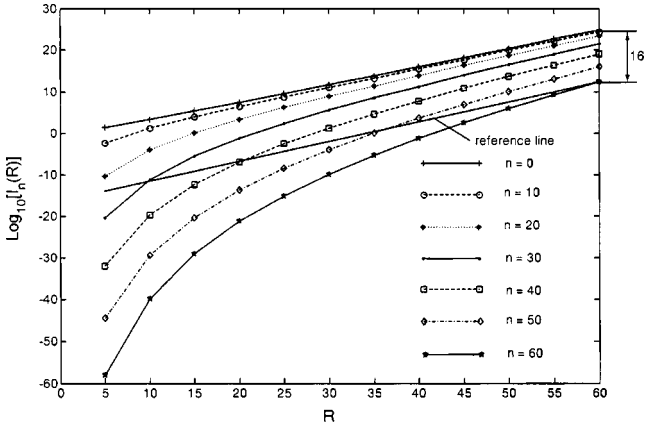


FIG. 2. Magnitude of the modified Bessel function of the first kind as functions of R and n.

$\geqslant n_{\min}$ have negligible contribution to the final value of the series. A list of R and $n_{\min}$ values and their corresponding $I_0(R)$, $I_1(R)$, $I_0(R)/I_{n_{\min}}(R)$ and $I_1(R)/I_{n_{\min}}(R)$ are shown in Table I. To further explain the "dynamic range" and $n_{\min}$, Fig. 2 is used to illustrate the magnitude of the modified Bessel function of the first kind as functions of R and n. With a reference line defined by $\log_{10}^{I_{\text{ref}}(R)/I_0(R)}=16$, any $n_{\min}$ value with a satisfactory approximation can be obtained by an interpolation technique when R ranges from 5 to 60 as shown in Fig. 2. For example, for $R=12.5$, $n_{\min}\approx 33$. For a large R, the corresponding coefficients to the first few expansion terms are also large. For this case, large number terms of the expansion are required to reduce the truncation error. However, the extra terms needed for accurate prediction may be outside the dynamic range defined by the machine accuracy of the computer. It is possible that the round-off errors of the subtractions of the first few terms are greater than the term around $n=n_{\min}$. For example, when $n=0$ and $R=5$, $I_0(5)=2.72\times 10^1$, the round-off error is of order $O(10^{-15})$; when $n=0$ and $R=50$, $I_0(50)=2.93\times 10^{20}$, the order of its round-off error increases to $O(10^4)$. As a result large prediction error occurs.

To overcome this difficulty, a small R is used in the AWP expansion. A splitting technique is introduced to replace R in the acoustical wave propagator in Eq. (8) with multiple components. Equation (8) can be rewritten as

$$e^{-R\hat{H}'}=e^{-\Sigma_{i=1}^{N_R}R_i'\hat{H}'}, \tag{23}$$

where $R_i'=R/N_R$, such that each propagator $e^{-R_i'\hat{H}'}$ can be calculated with sufficient prediction accuracy when the Chebyshev polynomial expansion is used to implement the operation of the acoustical wave propagator.

IV. RESULTS AND DISCUSSION

To demonstrate the implementation of the AWP on an initial wave packet in a thin plate, we select the following initial state vector:

$$\phi(x,y,0)=\begin{bmatrix}0\\ M_x(x,y,0)\\ M_y(x,y,0)\\ M_{xy}(x,y,0)\end{bmatrix}, \tag{24}$$

where $M_x(x,y,0)$, $M_y(x,y,0)$, and $M_{xy}(x,y,0)$ are related to the initial displacement $W(x,y,0)=W_0e^{-(x^2+y^2)/4\sigma^2}$. In this simulation, parameters of the steel plate used for numerical calculation are given as follows: $E=21.6\times 10^{10}\,\text{N/m}^2$, $\upsilon=0.3$, $2h=0.002\,\text{m}$, $\rho=7800\,\text{kg/m}^3$, $W_0=0.001\,\text{m}$, $\sigma=0.1$, the grid size N is 100, and the spatial sampling interval $\Delta x=\Delta y$ is 0.1 m. A plate with infinite length in the x and y dimensions is used, and the evolution of the wave packet is observed in the spatial range ($-5\,\text{m}\leqslant x\leqslant 5\,\text{m}$, $-5\,\text{m}\leqslant y\leqslant 5\,\text{m}$).

TABLE I. Dynamic range of the expansion coefficients of Chebyshev polynomials.

R	$n_{\min}$	$I_0(R)$	$I_1(R)$	$I_0(R)/I_{n_{\min}}(R)$	$I_1(R)/I_{n_{\min}}(R)$
5	25	2.72×10^1	2.43×10^1	3.74×10^{16}	3.35×10^{16}
10	32	2.82×10^3	2.67×10^3	1.50×10^{16}	1.43×10^{16}
15	38	3.40×10^5	3.28×10^5	2.41×10^{16}	2.33×10^{16}
20	43	4.36×10^7	4.25×10^7	2.86×10^{16}	2.79×10^{16}
25	47	5.77×10^9	5.66×10^9	1.77×10^{16}	1.74×10^{16}
30	51	7.82×10^{11}	7.69×10^{11}	1.97×10^{16}	1.93×10^{16}
35	55	1.07×10^{14}	1.06×10^{14}	3.12×10^{16}	3.07×10^{16}
40	58	1.49×10^{16}	1.47×10^{16}	1.93×10^{16}	1.91×10^{16}
45	61	2.08×10^{18}	2.06×10^{18}	1.56×10^{16}	1.54×10^{16}
50	64	2.93×10^{20}	2.90×10^{20}	1.53×10^{16}	1.51×10^{16}
55	67	4.15×10^{22}	4.11×10^{22}	1.74×10^{16}	1.72×10^{16}
60	70	5.89×10^{24}	5.85×10^{24}	2.21×10^{16}	2.19×10^{16}

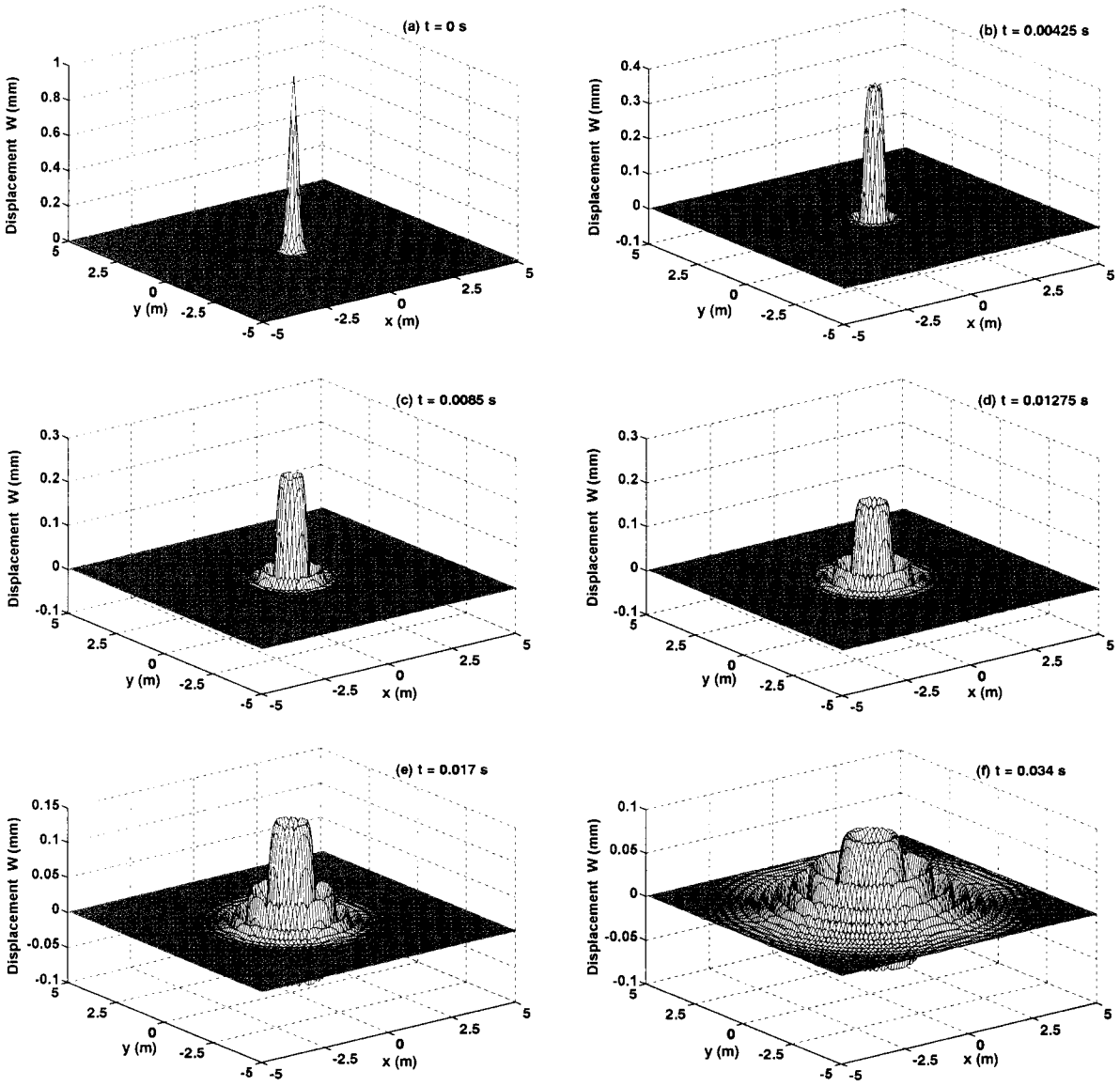


FIG. 3. Distributed flexural wave displacement W (mm) at different times.

Both CPE/FFT and Euler schemes are used to implement the AWP operation on the selected initial state vector. The results are assessed by comparison with the exact analytical solution [Eqs. (18) and (21)]. Figure 3 shows the time evolution of the displacement in the plate by the CPE/FFT scheme at 6 time steps. An initial Gaussian displacement distribution is illustrated in Fig. 3(a). At $t=0.004\,25$ s, we observe a little spread in the distribution of the displacement including negative components [see Fig. 3(b)]. At $t=0.0085$ s, the distributed flexural wave displacement has two principal crests and continues to spread out with decreased magnitude. After $t=0.012\,75$ s, we note the formation of subsidiary ''ripples'' ahead of the principal crest of each wave. This is due to the dispersive feature of the flexural wave in the plate. The high-frequency (short-wavelength) components in the wave packet spread outward faster than the lower frequency ones. The crests and troughs result from the ''superposition'' of different frequency components in the wave packet. Due to the frequency dependence of the traveling speed, the intensified (positive for crests; negative for troughs) locations also vary with time. Based on the wave speed of the component with the highest frequency in the packet, the flexural wave displacement distributes the whole observation range of the plate at $t=0.034$ s.

Figure 4 shows the time evolution of the moment in the plate by the CPE/FFT scheme corresponding to the displacements in Fig. 3. An initial moment distribution is illustrated in Fig. 4(a) and includes negative components. At $t=0.004\,25$ s, we observe the second crest in the distributed flexural wave moment. Compared with Fig. 4(a), the magnitude of the moment decreases quickly. It is noted that, when $t=0.012\,75$ s, the magnitude of the second principal crest is larger than that of the first one closer to the center. This is because the largest moment corresponds to the largest curvature of the displacement, rather than the largest magnitude. As the example used here is axisymmetric, displacement and moment distributions in the radial direction give rise to complete information about the wave packet. Based on the expression of the bending wave velocity of a thin plate, we

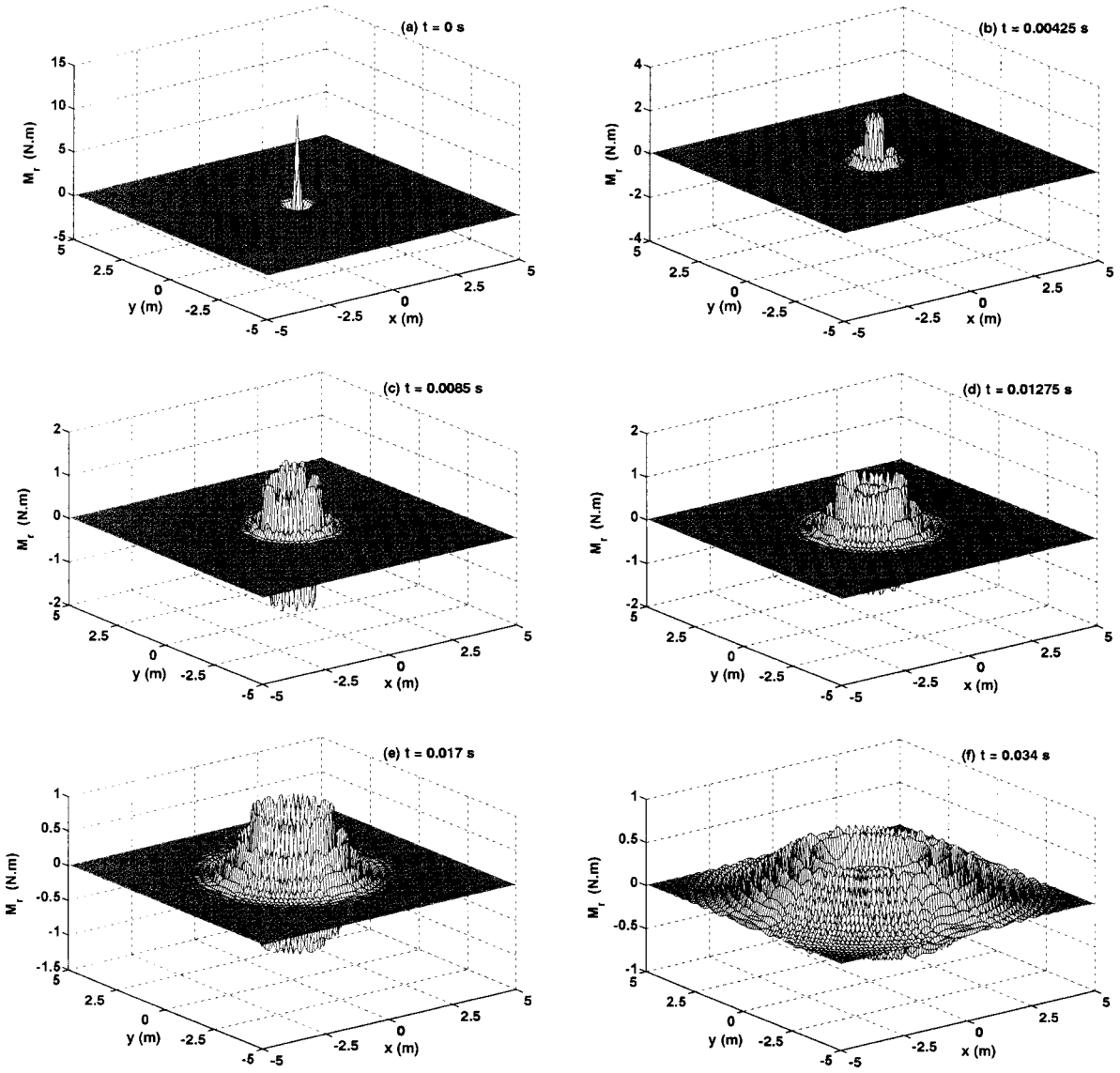


FIG. 4. Distributed flexural wave moment M_r (Nm) at different times.

note that different frequency components of bending waves travel at different wave speeds and they are therefore dispersive. As time increases, the effect of this dispersion on the distribution becomes obvious as shown in Fig. 4.

For the Chebyshev scheme, the time step is $dt_C = 0.0003$ s. Sixty terms of Chebyshev polynomials are used in this expansion. Total $N_R = 113$ time steps are used successively to give the displacement at $t = 0.034$ s. For the Euler scheme, the time step dt_E is 0.000 0001 s. The overall error of displacement is defined as

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(W_{\text{ex}} - W_{\text{pr}})^2}, \tag{25}$$

where W_{ex} and W_{pr} represent the exact displacement and predicted displacement by either the Chebyshev or Euler scheme at $t = 0.034$ s. Table II shows the maximum absolute error of displacement between the predictions of the Chebyshev scheme and the exact solution, and the Euler scheme and the exact solution when $y = 0$ m and $t = 0.034$ s.

Similarly, the overall error of moment between the predictions of the Chebyshev/Euler scheme and the exact solution is defined as

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_{\text{ex}} - M_{\text{pr}})^2}, \tag{26}$$

where M_{ex} and M_{pr} represent the exact bending moment and predicted moment by either the Chebyshev or Euler scheme. Table II shows the maximum absolute error of moment between the predictions of the Chebyshev scheme and the exact solution, and the Euler scheme and the exact solution when $y = 0$ m and $t = 0.034$ s. Further decrease in the time step will slowly reduce the overall error, both in the displacement and the moment. The finite-difference schemes based on the Taylor expansion of the time-evolution operator, even to high order, require extremely small time steps to ensure their convergence. In contrast, the Chebyshev scheme allows larger time steps. Large time steps for wave propagators have been proved to be possible for the Schrödinger equation, as the expansion coefficients are Bessel functions of the first kind.

TABLE II. The maximum overall errors of displacements and moments.

The maximum displacement error	Between the Chebyshev scheme and the exact solution	3.414×10^{-7}
	Between the Euler scheme and the exact solution	2.827×10^{-3}
The maximum moment error	Between the Chebyshev scheme and the exact solution	4.478×10^{-4}
	Between the Euler scheme and the exact solution	4.65×10^{-2}

However, expansion coefficients for the AWP are the modified Bessel functions of the first kind, which have a dynamic limit that prohibits the use of large time steps. Progress has been made to expand the AWP with coefficients of Bessel functions of the first kind, by defining modified Chebyshev polynomials. The publication of the new theoretical results is under way.[20]

The fast Fourier transformation (FFT) algorithm is the most time-consuming part of the AWP operation. To describe the numerical efficiency of the Chebyshev and Euler schemes, the numbers of calls of the FFT routine are estimated by Eqs. (27) and (28):[13]

$$N_{\mathrm{FFT}}^{E}=\frac{4t_E}{dt_E}, \tag{27}$$

$$N_{\mathrm{FFT}}^{C}=\frac{4t_C}{dt_C}N_C, \tag{28}$$

where t_E and dt_E represent the total time and the time step of wave evolution in the Euler scheme, respectively; dt_C, t_C and N_C represent the time step, the total time and the number of Chebyshev polynomials used in the expansion, respectively. In this paper, for the Euler scheme, $N_{\mathrm{FFT}}^{E}=1.36\times10^{6}$, and the overall errors of the displacement and bending moment are 9.718×10^{-5} and 1.157×10^{-2}, respectively. For the Chebyshev scheme, $N_{\mathrm{FFT}}^{C}=2.72\times10^{4}$, and the overall errors of the displacement and bending moment are 4.499×10^{-8} and 7.198×10^{-4}, respectively. This comparison shows a significant increase of numerical efficiency and accuracy when the Chebyshev polynomial scheme is used.

V. CONCLUSIONS

In this paper the acoustical wave propagator technique is extended to describe the time-domain evolution of wave packets in two-dimensional thin plates. A Chebyshev polynomial expansion scheme is implemented to carry out the operation of the acoustical wave propagator. Compared with the exact analytical solution, this scheme is found to be accurate and computationally effective in predicting the acoustical wave propagation in thin plates. This technique has been used for the preliminary study of dynamic stress concentration at structural discontinuities in a stepped plate[21] and a heterogeneous plate with multiple cylindrical patches.[22]

ACKNOWLEDGMENTS

The authors are grateful to Dr. J. B. Wang and Dr. S. Wang for valuable discussion, and their help in numerical analysis. S. Z. Peng thanks the University of Western Australia for financial support from the Dr. Julian Hunka Scholarship and the Australian Postgraduate Research Scholarship.

[1] D. Botteldooren, "Finite-difference time-domain simulation of low frequency room acoustic problems," J. Acoust. Soc. Am. **98**, 3302–3308 (1995).
[2] J. B. Schmeider and O. M. Ramahi, "The complementary operators method applied to acoustic finite-difference time-domain simulations," J. Acoust. Soc. Am. **104**, 686–693 (1998).
[3] L. Sbardella, "A time-domain method for the prediction of sound attenuation in lined ducts," J. Sound Vib. **239**, 379–396 (2001).
[4] N. Balakrishnan, C. Kalyanraman, and N. Sathyamurthy, "Time-dependent quantum mechanical approach to reactive scattering and related processes," Phys. Rev. **280**, 79–144 (1997).
[5] C. Leforestier, "A comparison of different propagation schemes for the time domain dependent Schrödinger equation," J. Comput. Phys. **94**, 59–80 (1991).
[6] K. Y. Fung, "Time-domain computation of acoustics in confinements," Proceedings of the 5th International Congress on sound and vibration, 15–18 Dec. 1997, pp. 1839–1847.
[7] A. Asker and A. S. Cakmak, "Explicit integration method for the time-dependent Schrödinger equation for collision problems," J. Chem. Phys. **68**, 2794–2798 (1997).
[8] M. D. Feit, J. A. Flect, and A. Steigeer, "Solution of the Schrödinger equation by a spectral method," J. Comput. Phys. **47**, 412–433 (1982).
[9] H. Tal-Ezer and R. Kosloff, "An accurate and efficient scheme for propagating the time-dependent Schrödinger equation," J. Chem. Phys. **81**, 3967–3971 (1984).
[10] P. Olsson, "Wave splitting of the Timoshenko beam equation in the time domain," Z. Angew. Math. Phys. **45**, 866–881 (1994).
[11] A. Karlsson, "Wave propagators for transient waves in one-dimensional media," Wave Motion **24**, 85–99 (1996).
[12] D. V. J. Billger and P. D. Folkow, "Wave propagators for the Timoshenko beam," Wave Motion **1122**, 1–20 (2002).
[13] J. Pan and J. B. Wang, "Acoustical wave propagator," J. Acoust. Soc. Am. **108**, 481–487 (2000).
[14] Y. Y. Yu, *Vibrations of Elastic Plates* (Springer, New York, 1996).
[15] M. P. Norton, *Fundamentals of Noise and Vibration Analysis for Engineers* (Syndicate Press, Cambridge, 1989).
[16] J. Halliwell and A. Zoupas, "Post-decoherence density matrix propagator for quantum Brownian motion," Phys. Rev. D **55**, 4697–4704 (1997).
[17] J. B. Wang, "Numerical differentiation using Fourier," The Mathematica Journal **8**, 1–5 (2002).
[18] K. F. Graff, *Wave Motion in Elastic Solids* (Clarendon, Oxford, 1975).
[19] T. J. Rivlin, *Chebyshev Polynomial: From Approximation Theory to Algebra and Number Theory* (Wiley, New York, 1990).
[20] J. B. Wang and J. Pan, A note on expansion of acoustical wave propagator using Chebyshev polynomials.
[21] S. Z. Peng and J. Pan, "Acoustical wave propagator technique for time-domain analysis of dynamic stress in a stepped plate," Proceedings of 2003 SAE Noise and Vibration Conference, Michigan, USA, 2003, NVC12.
[22] S. Z. Peng, J. Pan, and K. S. Sum, "Acoustical wave propagator method for time-domain analysis of flexural wave scattering and dynamic stress concentration in a heterogeneous plate with multiple cylindrical patches," Proceedings of the Tenth International Congress on sound and vibration, Stockholm, Sweden, 2003, pp. 5021–5028.

A simple formula for the longitudinal modes in a cylinder

Faiz Ahmad[a)]
Department of Mathematics, College of Science, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia

(Received 6 May 2003; revised 3 October 2003; accepted 17 November 2003)

For $c_T<c<c_L$, the Pochhammer dispersion relation can be replaced by an approximate equation which can be solved to obtain the dimensionless wave number *explicitly* as a function of the phase velocity. This equation is similar to the corresponding equation for the Lamb modes in plate. The two types of modes are interlaced, i.e., between every pair of modes of the plate there is a longitudinal mode of cylinder and vice versa. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1639907]

PACS numbers: 43.20.Bi, 43.20.Mv, 43.20.Ks [ANN] Pages: 475–477

I. INTRODUCTION

Consider an infinite cylinder of radius a. The cylinder is of isotropic material characterized by the phase speeds c_T and c_L, respectively, of the transverse and longitudinal bulk waves. Let ω and k denote the frequency and the wave number of a wave which propagates in a direction parallel to the axis of the cylinder whose surface is assumed to be free of traction. The dispersion relation for such a *longitudinal mode* is given by [see Ref. 1, Chap. 6],

$$\frac{2p}{a}(q^2+k^2)J_1(pa)J_1(qa)-(q^2-k^2)^2J_0(pa)J_1(qa) -4k^2pqJ_1(pa)J_0(qa)=0, \tag{1}$$

where

$$p=\sqrt{\frac{\omega^2}{c_L^2}-k^2}, \tag{2}$$

$$q=\sqrt{\frac{\omega^2}{c_r^2}-k^2}, \tag{3}$$

and J_0 and J_1 denote Bessel functions of the first kind. Equation (1) is known as the Pochhammer frequency equation. If k is fixed and Eq. (1) is solved for ω, we get an infinite number of frequencies for each wave number. Dispersion curves expressing ω, or the phase speed, as a function of the *dimensionless wave number ak*, give the *spectrum* of the possible longitudinal modes in the cylinder. The sketching of modes in a plate as well as a cylinder was made relatively easy by the work of Mindlin and his co-workers.[2,3]

If we consider wave propagation in an infinite plate of thickness $2h$, the dispersion relation for the *symmetric* modes is given by [Ref. 1, Chap. 6]

$$\frac{\tan(qh)}{\tan(ph)}=-\frac{4pqk^2}{(q^2-k^2)^2}. \tag{4}$$

Although the two dispersion relations (1) and (4) look very different from each other, their spectra will come out to be rather similar to each other.

a)Electronic mail: faizmath@hotmail.com

Recently, it was shown by the present author[4] that for $c_T<c<c_L$, Eq. (4) can be replaced by a simpler approximate equation which admits exact solutions. In the present paper we shall show that a similar procedure can be adopted for the dispersion relation (1) and an approximate solution can be obtained which expresses the wave number *explicitly* as a function of the phase speed of a mode. The *form* of both approximate solutions turns out to be the same. This explains why the qualitative features of the two spectra happen to be identical. The similarity becomes even more accentuated when we plot the two types of modes in the same figure. The cylinder modes are of the same shape as those of the plate, only they are displaced to the right, at each point, by $\pi/(4\sqrt{1-c^2/c_T^2})$ units.

II. THE APPROXIMATE EQUATION

We denote the phase speed and the dimensionless wave number of a mode having frequency ω and wave number k, respectively, by c and u, i.e.,

$$c=\omega/k \quad \text{and} \quad u=ak.$$

We shall only consider $c_T<c<c_L$. Let

$$\alpha(c)=\sqrt{\frac{c^2}{c_T^2}-1}, \tag{5}$$

$$\beta(c)=\sqrt{1-\frac{c^2}{c_L^2}}, \tag{6}$$

$$\gamma(c)=\frac{4\alpha(c)\beta(c)}{(1-\alpha^2(c))^2}. \tag{7}$$

We shall usually suppress the dependence of α, β, and γ on c. With the above definitions, Eq. (1) can be written as

$$\frac{I_0(u\beta)J_1(u\alpha)}{I_1(u\beta)J_0(u\alpha)}+\frac{2c^2\beta}{c_T^2(1-\alpha^2)^2u}\frac{J_1(u\alpha)}{J_0(u\alpha)}-\gamma=0, \tag{8}$$

where I_0 and I_1 denote the modified Bessel functions of the first kind. Suppose u, $u\alpha$, and $u\beta$ are all large compared with unity. To first order in $1/u$, we have the following asymptotic expression:[5]

$$\frac{I_0(u\beta)}{I_1(u\beta)} \sim 1 + \frac{1}{2\beta u}. \tag{9}$$

Substitution of Eq. (9) in Eq. (8) gives

$$\frac{J_1(u\alpha)}{J_0(u\alpha)}\left(1 + \frac{A}{u\alpha}\right) = \gamma, \tag{10}$$

where

$$A = \left(\frac{1}{2\beta} + \frac{2c^2\beta}{c_T^2(\alpha^2-1)^2}\right)\alpha.$$

Now, to first order in $1/u$, we have the following asymptotic expressions:[5]

$$J_0(u\alpha) \sim \sqrt{\frac{2}{\pi u\alpha}}\left[\cos\left(u\alpha - \frac{\pi}{4}\right) + \frac{1}{8u\alpha}\sin\left(u\alpha - \frac{\pi}{4}\right)\right],$$

$$J_1(u\alpha) \sim \sqrt{\frac{2}{\pi\alpha u}}\left[\sin\left(u\alpha - \frac{\pi}{4}\right) + \frac{3}{8u\alpha}\cos\left(u\alpha - \frac{\pi}{4}\right)\right].$$

Using the above expressions in Eq. (10), we obtain

$$\frac{\tan\left(u\alpha - \frac{\pi}{4}\right) + \frac{3}{8u\alpha}}{1 + \frac{1}{8u\alpha}\tan\left(u\alpha - \frac{\pi}{4}\right)}\left(1 + \frac{A}{u\alpha}\right) = \gamma. \tag{11}$$

From Eq. (11) we get, to first order in $1/u\alpha$

$$\tan\left(u\alpha - \frac{\pi}{4}\right) = \frac{\gamma - \frac{3}{8u\alpha}}{1 + \frac{A}{u\alpha} - \frac{\gamma}{8u\alpha}}. \tag{12}$$

It is worth mentioning that those values of $u\alpha$ which make $\tan(u\alpha - [\pi/4])$ large, occur away from the roots of Eq. (11); therefore, we may ignore them in the transition to Eq. (12). The last equation has been obtained under the assumption $u\alpha \gg 1$; hence, to a first approximation we drop all terms having $u\alpha$ in the denominator and obtain

$$\tan\left(u\alpha - \frac{\pi}{4}\right) = \gamma. \tag{13}$$

Equation (13) has the solution

$$u_n^0 = \frac{1}{\alpha}\left(n\pi + \frac{\pi}{4} + \tan^{-1}\gamma\right). \tag{14}$$

The above solution can be improved by substituting u_n^0 for u on the right side of the expression in Eq. (12). We thus find

$$u_n = \frac{1}{\alpha}\left(n\pi + \frac{\pi}{4} + \tan^{-1}\eta\right), \tag{15}$$

where

$$\eta = \frac{\gamma - \frac{3}{8\alpha u_n^0}}{1 + \frac{A}{\alpha u_n^0} - \frac{\gamma}{8\alpha u_n^0}}. \tag{16}$$

TABLE I. Dimensionless wave number for various longitudinal modes of an aluminum cylinder, $v = 0.355$, corresponding to the phase velocity $c = 1.8c_T$.

Modes	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Approximate Eq. (15)	3.162	5.331	7.465	9.585	11.698	13.806	15.913
Exact Eq. (1)	3.078	5.304	7.453	9.578	11.693	13.803	15.911

III. DISPERSION CURVES

Equations (14) and (15) are *explicit* solutions of the Pochhammer equation (1) and can be used to plot the dispersion curves corresponding to the modes which may be denoted by $s_1, s_2, \ldots$. The approximate results do not apply to the lowest mode s_0, which has $u = 0$ on the one end, where c equals the bar velocity, and $\alpha = 0$ on the other end where it crosses the line $c = c_T$. In Table I we compare the exact results for aluminum when $c = 1.8c_T$, with the approximate ones given by Eq. (15). We note that the error in s_1 is less than 3 parts in 100, while in s_2 it is nearly 5 parts in 1000. It steadily decreases as we move up the sequence.

In Fig. 1 we have plotted the longitudinal modes s_1, s_2, s_3, and s_4. The solid curves represent the exact solution, while the dashed ones represent the approximate solution obtained from Eq. (15). We find that the two curves for the modes s_2 onward match each other almost in the entire range

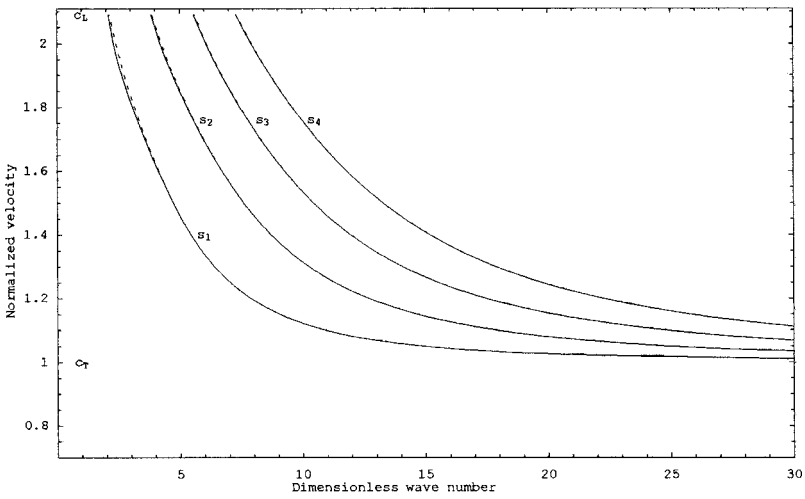


FIG. 1. Longitudinal modes $s_1, \ldots, s_4$ in an aluminum cylinder. Solid curves are exact and the dashed curves are based on Eq. (15).

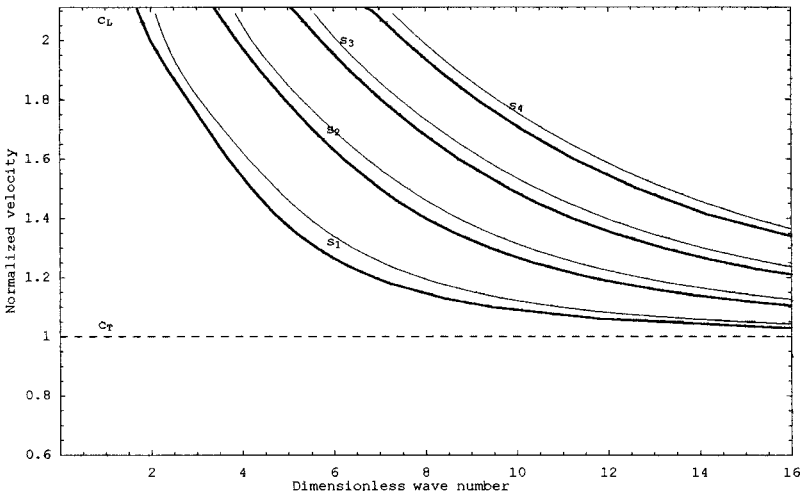


FIG. 2. Symmetric modes of a plate and the longitudinal modes of a cylinder shown together. The plate modes are shown in bold.

$c_T < c < c_L$. The curves for the first mode differ slightly between $1.5c_T < c < c_L$. The error in this mode at $1.5c_T$ is less than 3 parts in 1000; it increases to 4 percent at $c = 1.92c_T$ then decreases as c approaches the speed c_L. In this limit the exact equation becomes

$$J_1(u\alpha) = 0, \tag{17}$$

while the approximate Eq. (12) gives

$$\tan\left(u\alpha - \frac{\pi}{4}\right) = 0. \tag{18}$$

The first three zeros of J_1 occur at 0, 3.8317, and 7.0156, and Eq. (18) has the solutions

$$u_n = n\pi + \frac{\pi}{4}, \quad n = 0,1,2,\ldots. \tag{19}$$

The second zero of the Bessel function J_1 corresponds to the mode s_1 and the approximate value $5\pi/4$, given by Eq. (19), differs from it by less than 3 percent, while the error in the next value is less than 3 parts in 1000.

IV. COMPARISON WITH LAMB MODES

It was recently shown[4] that the dispersion relation (4) for the symmetric modes in a plate can be approximated by the equation

$$\nu_n = \frac{1}{\alpha}(n\pi + \tan^{-1}\gamma), \tag{20}$$

where $\nu = hk$ denotes the dimensionless wave number. A comparison of Eqs. (15) and (20) brings out the similarity of the two approximate equations with the result that the dispersion curves for a plate and a cylinder are identical in shape but the curve representing a longitudinal mode of a cylinder is displaced to the right of the corresponding curve for a plate, at each point, by $\pi/[4\alpha(c)]$ units. In Fig. 2 we display the modes for a plate as well as a cylinder in the same figure. The plate and cylinder modes s_1, s_2, etc. appear in pairs, the plate mode being on the left. Between every pair of consecutive modes of a plate there is a mode of the cylinder and vice versa. Along a horizontal line c = constant, consecutive modes of a plate are equally spaced. The same is true for the modes of the cylinder.

[1] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, Amsterdam, 1980).
[2] R. D. Mindlin, "Waves and vibrations in isotropic elastic plates," in *First Symposium on Structural Mechanics, 1958*, edited by J. N. Goodier and N. J. Hoff (Pergamon, Oxford, 1960), pp. 199–232.
[3] M. Onoe, H. D. McNiven, and R. D. Mindlin, "Dispersion of axially symmetric waves in elastic rods," J. Appl. Mech. **28**, 729–734 (1963).
[4] F. Ahmad, "A simple formula for the Lamb modes in a plate," J. Acoust. Soc. Am. **111**, 1974–1975 (2002).
[5] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, D.C., 1972).

Acoustic field induced in spheres with inhomogeneous density by external sources

Gerassimos C. Kokkorakis and John G. Fikioris
Institute of Communication and Computer Systems, National Technical University of Athens, GR-157 73 Zografou, Greece

(Received 22 March 2003; revised 29 September 2003; accepted 30 October 2003)

Acoustic or electromagnetic fields induced in the interior of inhomogeneous penetrable bodies by external sources can be evaluated via well-known volume integral equations. For bodies of arbitrary shape and/or composition, for which separation of variables fails, a direct attack for the solution of these integral equations is the only available approach. In a previous paper by the same authors the scalar (acoustic) field in inhomogeneous spheres of arbitrary compressibility, but with constant density, was considered. In the present one the direct hybrid (analytical-numerical) method applied to the much simpler integral equation for spheres with constant density is generalized to densities that vary with r, θ, or even φ. This extension is by no means trivial, owing to the appearance of the derivatives of both the density and the unknown function in the volume integral, a fact necessitating a more subtle and accuracy-sensitive approach. Again, the spherical shape allows use of the orthogonal spherical harmonics and of Dini's expansions of a general type for the radial functions. The convergence of the latter, shown to be superior to other possible sets of orthogonal expansions, can be further optimized by the proper selection of a crucial parameter in their eigenvalue equation.
 [DOI: 10.1121/1.1635410]

PACS numbers: 43.20.Fn [GCG] Pages: 478–487

I. INTRODUCTION

The motivation for solving the volume integral equation for penetrable spheres with varying density $\rho(\mathbf{r})$, the general mathematical difficulties of various approaches used in the past as well as the specific ones that are successfully faced by our direct hybrid approach, are discussed in a previous paper[1] and will not be repeated here. For convenience, we repeat here the general volume integral equation in acoustics[1–4]

$$\Phi(\mathbf{r})=\Phi^{\mathrm{inc}}(\mathbf{r})+\int_V k_0^2\left[\frac{b(\mathbf{r}')}{b_0}-1\right]\Phi(\mathbf{r}')g(\mathbf{r},\mathbf{r}')dV' + \int_V \nabla'\cdot\left\{\left[\frac{\rho_0}{\rho(\mathbf{r}')}-1\right]\nabla'\Phi(\mathbf{r}')\right\}g(\mathbf{r},\mathbf{r}')dV', \tag{1}$$

where $\rho(\mathbf{r})$, $b(\mathbf{r})$ are the varying density and compressibility, respectively, of the medium in V surrounded by a medium of constant ρ_0, b_0 values. An incident field $\Phi^{\mathrm{inc}}(\mathbf{r})$ impinges on V and induces the total (unknown) scalar field $\Phi(\mathbf{r})$. The harmonic time dependence is $\exp(i\omega t)$ and

$$g(\mathbf{r},\mathbf{r}')=e^{-ik_0R}/4\pi R,\quad R=|\mathbf{r}-\mathbf{r}'|,\quad k_0=\omega(b_0\rho_0)^{1/2}, \tag{2}$$

is the scalar Green's function of free space, which can be expanded in spherical coordinates as follows:[5]

$$g(\mathbf{r},\mathbf{r}')=\frac{e^{-ik_0R}}{4\pi R} =\frac{-ik_0}{4\pi}\sum_{n'=0}^{\infty}(2n'+1)j_{n'}(k_0r_<)h_{n'}(k_0r_>) \cdot\sum_{m'=-n'}^{n'}\frac{(n'-m')!}{(n'+m')!}P_{n'}^{m'}(\cos\theta)P_{n'}^{m'}(\cos\theta') \times e^{im'(\varphi-\varphi')}. \tag{3}$$

Here, $h_n\equiv h_n^{(2)}=j_n-in_n$ is the spherical Hankel function and $r_>(r_<)$ stands for the larger (smaller) of r, r'. Again, when $\mathbf{r}$ is exterior to V Eq. (1) provides a relation for the exterior (scattered) field; when $\mathbf{r}$ is inside V the equation stands as the volume integral equation for the evaluation of the total internal field $\Phi(\mathbf{r})$.

II. TRANSFORMATION OF THE INTEGRAL EQUATION

We first give an alternative form for the integrand of (1), apart from $g(\mathbf{r},\mathbf{r}')$

$$k_0^2\left[\frac{b(\mathbf{r})}{b_0}-1\right]\Phi(\mathbf{r})+\nabla\cdot\left\{\left[\frac{\rho_0}{\rho(\mathbf{r})}-1\right]\nabla\phi(\mathbf{r})\right\} =k_0^2\left[\frac{b(\mathbf{r})\rho(\mathbf{r})}{b_0\rho_0}-1\right]\Phi(\mathbf{r})+\nabla[\rho_0-\rho(\mathbf{r})]\cdot\frac{\nabla\Phi(\mathbf{r})}{\rho(\mathbf{r})}. \tag{4}$$

This second form, which avoids second derivatives of $\Phi(\mathbf{r})$, follows easily from the differential equation that $\Phi(\mathbf{r})$ satisfies in V, namely[2–4]

$$\nabla\cdot\left[\frac{\nabla\Phi(\mathbf{r})}{\rho(\mathbf{r})}\right]=-\omega^2b(\mathbf{r})\rho(\mathbf{r}). \tag{5}$$

Following now the method of analysis proposed in Ref. 6, for the expression of electromagnetic field quantities in spheres, we expand the scalar field quantities of (1) in spherical harmonics

$$k_0^2\left[\frac{b(\mathbf{r})\rho(\mathbf{r})}{b_0\rho_0}-1\right]\Phi(\mathbf{r})+\nabla[\rho_0-\rho(\mathbf{r})]\cdot\frac{\nabla\Phi(\mathbf{r})}{\rho(\mathbf{r})}=\sum_{n=0}^{\infty}\sum_{m=-n}^{n}g_{nm}(r)P_n^m(\cos\theta)e^{im\varphi},\tag{6}$$

$$\Phi(\mathbf{r})=\sum_n\sum_m f_{nm}(r)P_n^m(\cos\theta)e^{im\varphi},\tag{7}$$

$$\Phi^{\mathrm{inc}}(\mathbf{r})=\sum_n\sum_m d_{nm}(r)P_n^m(\cos\theta)e^{im\varphi},\tag{8}$$

where g, f, and d represent the radial part of the developments of the functions of $\mathbf{r}$ on the left-hand side; their calculation is the object of the present paper. As a starting point we further suppose that $\rho(\mathbf{r})$ is continuous in $0\leqslant r\leqslant a$, namely that $\rho(a-,\theta,\varphi)=\rho_0$; the significance of this initial restriction, which will be lifted later, will become apparent in a little while. Substituting now (3) and (6)–(8) into (1), using the orthogonality relations of $P_n^m(\cos\theta)e^{im\varphi}$ first to integrate with respect to θ', φ' and then to separate the equation, we end up[1,6,7] with the relation

$$f_{nm}(r)=d_{nm}(r)-ik_0[h_n(k_0r)I_{nm}(r)+j_n(k_0r)L_{nm}(r)],\tag{9}$$

where, as in Ref. 6, we define

$$I_{nm}(r)=\int_0^r g_{nm}(r')j_n(k_0r')r'^2\,dr',\qquad L_{nm}(r)=\int_r^a g_{nm}(r')h_n(k_0r')r'^2\,dr'.\tag{10}$$

Also, together with Eq. (5) for $\Phi(\mathbf{r})$, Eq. (4) obtains the form

$$\begin{aligned}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}g_{nm}(\mathbf{r})P_n^m(\cos\theta)e^{im\varphi}&=\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left\{k_0^2\left[\frac{b(\mathbf{r})\rho(\mathbf{r})}{b_0\rho_0}-1\right]f_{nm}(r)P_n^m(\cos\theta)e^{im\varphi}\right.\\&+\frac{1}{\rho(\mathbf{r})}\frac{\partial[\rho_0-\rho(\mathbf{r})]}{\partial r}\frac{\partial f_{nm}(r)}{\partial r}P_n^m(\cos\theta)e^{im\varphi}\\&+\frac{1}{r^2\rho(\mathbf{r})}\frac{\partial[\rho_0-\rho(\mathbf{r})]}{\partial\theta}f_{nm}(r)\cdot\frac{\partial P_n^m(\cos\theta)}{\partial\theta}e^{im\varphi}\\&\left.+\frac{im}{r^2\sin^2\theta\rho(\mathbf{r})}\frac{\partial[\rho_0-\rho(\mathbf{r})]}{\partial\varphi}f_{nm}(r)P_n^m(\cos\theta)e^{im\varphi}\right\}.\end{aligned}\tag{11}$$

III. INHOMOGENEITY INDEPENDENT OF θ, φ

As in Ref. 1, we start from the simpler case $\rho(\mathbf{r})=\rho(r)$, $b(\mathbf{r})=b(r)$, independent of θ, φ. We further suppose that $d_{nm}=d_n$; this implies that for an incident plane wave its direction of propagation is along the z axis, $\mathbf{k}_0=k_0\hat{\mathbf{z}}$, and that no φ dependence is introduced in the problem via $\Phi^{\mathrm{inc}}(\mathbf{r})$. Thus, the index m disappears, $f_{nm}=f_n$, $g_{nm}=g_n$, and Eq. (11) simplifies to

$$g_n(r)=k_0^2\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]f_n(r)-\frac{\rho'(r)}{\rho(r)}f_n'(r).\tag{12}$$

Differentiation of $f_n(r)$ in (9) yields the simple result

$$f_n'(r)=d_n'(r)-ik_0^2[h_n'(k_0r)I_n(r)+j_n'(k_0r)L_n(r)].\tag{13}$$

Substitution of (13) and (9) in (12) leads to the equation

$$\begin{aligned}g_n(r)=k_0^2\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]d_n(r)-ik_0^3\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]&\times[h_n(k_0r)I_n(r)+j_n(k_0r)L_n(r)]-\frac{\rho'(r)}{\rho(r)}d_n'(r)\\&+ik_0^2\frac{\rho'(r)}{\rho(r)}[h_n'(k_0r)I_n(r)+j_n'(k_0r)L_n(r)].\end{aligned}\tag{14}$$

It is remarkable to observe that, as in Ref. 6, no derivatives of the unknown function $g_n(r)$ appear in (13) or (14), a significant fact from the standpoint of convergence of the Dini series expansions for $f_n(r)$ and $g_n(r)$ that are introduced later on. This is a clear advantage of the analytical approach developed here over other possible analytical and/or numerical methods. For the latter, in particular, it is impossible to avoid the appearance of derivatives of the unknown function $\Phi(\mathbf{r})$ in (1) or (6); this leads to complicated relations between $\Phi(\mathbf{r})$ and $\nabla\Phi(\mathbf{r})$ and to worsening of the convergence of the whole procedure.

Before proceeding with the further treatment of (14) we develop the corresponding expression in the case of discontinuous $\rho(r)$ at $r=a$, i.e., when $\rho(a-)=\rho_1\neq\rho_0$. This step discontinuity in $\rho(r)$ implies a delta-function behavior of its derivative at $r=a$ and a consequent additional contribution to the last integral in (1) from this term. To take account of this discontinuity, we define the density as follows:

$$\rho(r)=\rho_c(r)+(\rho_1-\rho_0)u(a-r),\tag{15}$$

where $\rho_c(r)$ is the continuous part of $\rho(r)$, with $\rho_c(a)=\rho_0$, $\rho(a)=\rho_1$, and $u(x)$ the unit step function, $u(x)=0$ for $x<0$ and $u(x)=1$ for $0\leqslant x$. Then

$$\rho'(r)=\rho_c'(r)-(\rho_1-\rho_0)\delta(a-r).\tag{16}$$

In this φ-independent case Eq. (12) remains unchanged and substitution of $\rho'(r)$ from (16) in it yields

$$g_n(r)=k_0^2\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]f_n(r)-\frac{f_n'(r)}{\rho(r)}\times[\rho_c'(r)-(\rho_1-\rho_0)\delta(a-r)].\tag{17}$$

The delta function $\delta(a-r)$ in (17) modifies now the derivation of (9)–(10) from (1) as follows:

$$f_n(r)=d_n(r)-ik_0\int_0^a r'^2 g_n(r')j_n(k_0r_<)h_n(k_0r_>)dr'$$

$$=\begin{cases} d_n(r)-ik_0[h_n(k_0r)\tilde{I}_n(r)+j_n(k_0r)\tilde{L}_n(r)]-ik_0a^2j_n(k_0r)h_n(k_0a)\left[1-\dfrac{\rho_0}{\rho_1}\right]f_n'(a), & r\leqslant a, \quad (18a) \\ d_n(r)-ik_0h_n(k_0a)\tilde{I}_n(a)-ik_0a^2h_n(k_0r)j_n(k_0a)\left[1-\dfrac{\rho_0}{\rho_1}\right]f_n'(a), & r\geqslant a, \quad (18b) \end{cases}$$

where

$$\tilde{I}_n(r)=\int_0^r \tilde{g}_n(r')j_n(k_0r')r'^2\,dr',$$
$$\tilde{L}_n(r)=\int_r^a \tilde{g}_n(r')h_n(k_0r')r'^2\,dr', \quad (19)$$

and

$$\tilde{g}_n(r)=k_0^2\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]f_n(r)-\frac{\rho_c'(r)}{\rho(r)}f_n'(r). \quad (20)$$

As before, we now substitute in this last equation $f_n(r)$ from (18) and $f_n'(r)$ from the same equation after differentiation

$$f_n'(r)=d_n'(r)-ik_0^2[h_n'(k_0r)\tilde{I}_n(r)+j_n'(k_0r)\tilde{L}_n(r)]$$
$$-ik_0^2a^2j_n'(k_0r)h_n(k_0a)\left[1-\frac{\rho_0}{\rho_1}\right]f_n'(a). \quad (21)$$

We thus obtain the final form that Eq. (14) assumes in this case

$$\tilde{g}_n(r)=k_0^2\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]\{d_n(r)-ik_0[h_n(k_0r)\tilde{I}_n(r)$$
$$+j_n(k_0r)\tilde{L}_n(r)]\}-\frac{\rho_c'(r)}{\rho(r)}\{d_n'(r)$$
$$-ik_0^2[h_n'(k_0r)\tilde{I}_n(r)+j_n'(k_0r)\tilde{L}_n(r)]\}$$
$$-ik_0^2a^2h_n(k_0a)\left[1-\frac{\rho_0}{\rho_1}\right]f_n'(a)$$
$$\times\left\{j_n(k_0r)k_0\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]-j_n'(k_0r)\frac{\rho_c'(r)}{\rho(r)}\right\}, \quad (22)$$

in which $f_n'(a)$, following from (21) with $\tilde{L}_n(a)=0$, has the value

$$f_n'(a)=\frac{d_n'(a)-ik_0^2h_n'(k_0a)\tilde{I}_n(a)}{1+ik_0^2a^2j_n'(k_0a)h_n(k_0a)\left[1-\dfrac{\rho_0}{\rho_1}\right]}. \quad (23)$$

Let us observe that when $\rho_1=\rho_0$ [no discontinuity in $\rho(r)$] we have $\tilde{g}_n(r)=g_n(r)$, $\tilde{I}_n=I_n$, $\tilde{L}_n=L_n$, and (22) reduces to (14), as expected. Therefore (22) is the more general form of the equation for $\tilde{g}_n(r)$ in either case. The reader may be further convinced that (22) is the proper form of the equation for $\tilde{g}_n(r)$ in the case of discontinuity $\rho_1\neq\rho_0$, if the following two facts are taken into account: (1) The special case $\rho_c(r)=\rho_0$ or $\rho(r)=\rho_1$ $(0\leqslant r\leqslant a)$ may be treated by setting $\tilde{g}_n(r)=B_nj_n(k_1r)$. It is then easy to prove that the correct expression for B_n, following from separation of variables, is obtainable by a direct solution of (22), rather than (14), following steps identical to those applied in the Appendix of Ref. 1 for the special case $\rho(\mathbf{r})=\rho_0$. (2) In the more general case of inhomogeneity dependent on r and θ analyzed later in Sec. VI, it is shown that the—corresponding to (15), (16)—more general expressions (54), (55) are the correct ones through which the boundary conditions at $r=a$ are explicitly satisfied. This is shown—for any $b(r,\theta)$, $\rho(r,\theta)$—on the basis of Eqs. (73)–(76), at the end of Sec. VI.

Finally, let us observe that for an incident plane wave

$$d_n(r)=(-i)^n(2n+1)j_n(k_0r),$$
$$d_n'(r)=(-i)^n(2n+1)k_0j_n'(k_0r). \quad (24)$$

IV. DEVELOPMENT OF THE RADIAL FUNCTIONS IN DINI SERIES

Proceeding now as in Ref. 1, in connection with the general equation (22) we expand $\tilde{g}_n(r)$ in a generalized Dini series

$$\tilde{g}_n(r)=\sum_{\mu=1}^{\infty} G_{n\mu}j_n\left(\gamma_{n\mu}\frac{r}{a}\right). \quad (25)$$

The spherical Bessel functions $j_n[\gamma_{n\mu}(r/a)]$ $(\mu=1,2,...)$ form a complete and orthogonal set of eigenfunctions in $0\leqslant r\leqslant a$[8,9] as long as the constants $\gamma_{n\mu}$ are chosen as the roots of the eigenvalue equation

$$\gamma_{n\mu}j_n'(\gamma_{n\mu})/j_n(\gamma_{n\mu})=-t_n(\mu=1,2,...). \quad (26)$$

Here, t_n is any chosen constant; as in Ref. 1, it will be chosen later so as to optimize the convergence of the Dini expansion for $f_n(r)$

$$f_n(r)=\sum_{\mu=1}^{\infty} B_{n\mu}j_n\left(\gamma_{n\mu}\frac{r}{a}\right). \quad (27)$$

We will need, also, the expansion of $j_n(k_0r)$ (Ref. 1)

$$j_n(k_0 r)=\sum_{\mu=1}^{\infty} J_{n\mu} j_n\left(\frac{\gamma_{n\mu}}{a} r\right), \tag{28}$$

$$J_{n\mu}=2\,\frac{\gamma_{n\mu}^2 j_n(k_0 a)\left[t_n+\dfrac{(k_0 a) j_n'(k_0 a)}{j_n(k_0 a)}\right]}{[\gamma_{n\mu}^2-(k_0 a)^2] j_n(\gamma_{n\mu})[\gamma_{n\mu}^2-(n-t_n+1)(n+t_n)]}$$

We next substitute $\tilde{g}_n(r')$ from (25) in the integrals for $\tilde{I}_n(r)$ and $\tilde{L}_n(r)$ of (19); following the steps of Ref. 1 we can integrate on the basis of standard integral formulas for the product of any two spherical Bessel functions of the same order;[1] the end result is

$$h_n(k_0 r)\tilde{I}_n(r)+j_n(k_0 r)\tilde{L}_n(r)$$
$$=\frac{ia^2}{k_0}\sum_{\mu=1}^{\infty}\frac{G_{n\mu}}{\gamma_{n\mu}^2-k_0^2 a^2}\times\left[j_n\left(\gamma_{n\mu}\frac{r}{a}\right)-i(k_0 a)^2 j_n(k_0 r) C_{n\mu}\right], \tag{29}$$

$$h_n'(k_0 r)\tilde{I}_n(r)+j_n'(k_0 r)\tilde{L}_n(r)$$
$$=\frac{ia}{k_0^2}\sum_{\mu=1}^{\infty}\frac{G_{n\mu}}{\gamma_{n\mu}^2-k_0^2 a^2}\times\left[\gamma_{n\mu} j_n'\left(\gamma_{n\mu}\frac{r}{a}\right)-i(k_0 a)^3 j_n'(k_0 r) C_{n\mu}\right], \tag{30}$$

$$C_{n\mu}=\frac{\gamma_{n\mu}}{k_0 a}h_n(k_0 a) j_{n+1}(\gamma_{n\mu})-h_{n+1}(k_0 a) j_n(\gamma_{n\mu})$$
$$=j_n(\gamma_{n\mu})\left[h_n'(k_0 a)+t_n\frac{h_n(k_0 a)}{k_0 a}\right]\quad n=0,1,...,\ \mu=1,2,...\,. \tag{31}$$

Referring now to (22), we observe that, in order to apply the orthogonal properties of eigenfunctions $j_n[\gamma_{n\mu}(r/a)]$ ($\mu=1,2,...$), we should first evaluate the following set of expansion coefficients:

$$Q_1(n,\mu)=\int_0^a j_n(k_0 r)\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right] j_n\left(\gamma_{n\mu}\frac{r}{a}\right) r^2\,dr$$
$$=\sum_{s=1}^{\infty} J_{ns} Q_2(n|\mu,s), \tag{32}$$

$$Q_2(n|\mu,\mu')=\int_0^a j_n\left(\gamma_{n\mu}\frac{r}{a}\right)\left[\frac{b(r)\rho(r)}{b_0\rho_0}-1\right]\times j_n\left(\gamma_{n\mu'}\frac{r}{a}\right) r^2\,dr, \tag{33}$$

$$Q_3(n,\mu)=\int_0^a j_n'(k_0 r)\,\frac{\rho_c'(r)}{\rho(r)}\, j_n\left(\gamma_{n\mu}\frac{r}{a}\right) r^2\,dr, \tag{34}$$

$$Q_4(n|\mu,\mu')=\int_0^a j_n\left(\gamma_{n\mu}\frac{r}{a}\right)\frac{\rho_c'(r)}{\rho(r)} j_n\left(\gamma_{n\mu'}\frac{r}{a}\right) r^2\,dr, \tag{35}$$

$$N_{n\mu}=\int_0^a j_n^2\left(\frac{\gamma_{n\mu}}{a} r\right) r^2\,dr$$
$$=\frac{a^3}{2}[j_n^2(\gamma_{n\mu})-j_{n-1}(\gamma_{n\mu}) j_{n+1}(\gamma_{n\mu})]$$
$$=\frac{a^3}{2} j_n^2(\gamma_{n\mu})\left[1-\frac{(n+1-t_n)(n+t_n)}{\gamma_{n\mu}^2}\right]. \tag{36}$$

The expression for $N_{n\mu}$ can be found in Ref. 1; also, $Q_2(n|\mu,\mu')=Q_2(n|\mu',\mu)$ but $Q_4(n|\mu,\mu')\neq Q_4(n|\mu',\mu)$. The above coefficients Q_1 to Q_4 can, in general, be evaluated by numerical integration. For an incident plane wave the expansions (24) for $d_n(r)$, $d_n'(r)$ are now substituted in (18) which, based on orthogonality, assumes the form

$$B_{n\mu}=\Delta_n J_{n\mu}+\frac{a^2 G_{n\mu}}{\gamma_{n\mu}^2-k_0^2 a^2},\quad n=0,1,...,\quad \mu=1,2,....;$$
$$\Delta_n=(-i)^n(2n+1)+ik_0 a^2 h_n(k_0 a)\left(\frac{\rho_0}{\rho_1}-1\right) f_n'(a)$$
$$-ik_0^2 a^4 T_n, \tag{37}$$

where the constant T_n is defined in terms of the $G_{n\mu}$ by the series

$$T_n=\sum_{q=1}^{\infty}\frac{C_{nq} G_{nq}}{\gamma_{nq}^2-k_0^2 a^2}, \tag{38}$$

while $f_n'(a)$ follows from (23). Observing that

$$\tilde{I}_n(a)=\int_0^a \tilde{g}_n(r) j_n(k_0 r) r^2\,dr$$
$$=\sum_{\mu=1}^{\infty}\int_0^a G_{n\mu} j_n(k_0 r) j_n\left(\gamma_{n\mu}\frac{r}{a}\right) r^2\,dr$$
$$=\sum_{\mu=1}^{\infty} J_{n\mu} N_{n\mu} G_{n\mu}, \tag{39}$$

we finally get

$$f_n'(a)=\frac{(-i)^n(2n+1)k_0 j_n'(k_0 a)-ik_0^2 h_n'(k_0 a)\Sigma_{p=1}^{\infty} J_{np} N_{np} G_{np}}{1-ik_0^2 a^2 j_n'(k_0 a) h_n(k_0 a)\left[\dfrac{\rho_0}{\rho_1}-1\right]}. \tag{40}$$

It is useful, also, to observe that substitution of $J_{n\mu}$, $N_{n\mu}$ from (28), (36) in (39) yields

$$T_n=\frac{\tilde{I}_n(a)}{k_0 a^4}\cdot\frac{h_n(k_0 a)t_n+k_0 a h_n'(k_0 a)}{j_n(k_0 a)t_n+k_0 a j_n'(k_0 a)}. \tag{41}$$

Finally, substitution of $B_{n\mu}$ from (37), $d_n(r)$ and $d_n'(r)$ from (24), $\tilde{g}_n(r)$ from (23), and the expressions (29) and (30), all in (22), and the use of the orthogonality of $j_n[\gamma_{n\mu}(r/a)]$ in $0\leqslant r\leqslant a$ leads to

$$N_{n\mu}G_{n\mu}=k_0\Delta_n[k_0Q_1(n|\mu)-Q_3(n|\mu)]+\sum_{q=1}^{\infty}\frac{G_{nq}}{\gamma_{nq}^2-k_0^2a^2}[k_0^2a^2Q_2(n|q,\mu)-a\gamma_{nq}Q_4(n|q,\mu)],\quad n=0,1,...,\quad \mu=1,2,.... \tag{42}$$

If T_n from (38) and $f_n'(a)$ from (40) are substituted in (42), we obtain a linear system of equations from which the $G_{n\mu}$'s are evaluated after truncation. The $B_{n\mu}$'s, for $f_n(r)$ then follow from (37).

V. CONVERGENCE AND OPTIMUM FORM OF DINI'S EXPANSIONS

As explained in Ref. 1 and 8, the Dini's series

$$g(r)=\sum_{\mu=1}^{\infty}A_{n\mu}j_n\left(\gamma_{n\mu}\frac{r}{a}\right),\quad 0\leq r\leq a \tag{43}$$

of a function $g(r)$, continuous in the indicated interval converges at least as far as $1/\mu^2$ as $\mu\to\infty$. This rate can be further improved if the particular choice

$$t_{\text{optimum}}=-\frac{ag'(a)}{g(a)}, \tag{44}$$

in the eigenvalue equation (27) is made.[1,8] With this particular choice the Dini expansion (43) converges at least as fast as $1/\mu^3$ and it may safely be differentiated.[1,8]

As further explained in Ref. 1, $j_n(x)\sim[\sin(x-n\pi/2)]/x$ as $x\to\infty$, while the roots $\gamma_{n\mu}\to\mu\pi$ and, from (36), $N_{n\mu}\sim 1/\mu^2$ as $\mu\to\infty$. With this in mind we can deduce from (29) that $J_{n\mu}\sim 1/\mu$, from (32) that $C_{n\mu}\sim 1/\mu$; from (38) that the series for T_n converges as $1/\mu^4$, and from (40) that the series for $f_n'(a)$ as $1/\mu^4$. All this shows that $B_{n\mu}$ from (37), due to the factor $J_{n\mu}$, converges as $1/\mu$ as $\mu\to\infty$ for arbitrary t_n. If, however, the optimum value for t_n is chosen in connection with the expansion (27) for $f_n(r)$, i.e., when

$$t_{n,\text{optimum}}=-\frac{af_n'(a)}{f_n(a)}\equiv\tilde{t}_n, \tag{45}$$

the sequence of $B_{n\mu}$ and the series for $f_n(r)$ will converge at least as $1/\mu^2$ and $1/\mu^3$, respectively.

Indeed, setting $r=a$ in (18) and (21), we get

$$f_n(a)=(-i)^n(2n+1)j_n(k_0a)-ik_0h_n(k_0a)\tilde{I}_n(a)-ik_0a^2j_n(k_0a)h_n(k_0a)\left[1-\frac{\rho_0}{\rho_1}\right]f_n'(a), \tag{46}$$

$$f_n'(a)=(-i)^n(2n+1)k_0j_n'(k_0a)-ik_0^2h_n'(k_0a)\tilde{I}_n(a)-ik_0^2a^2j_n'(k_0a)h_n(k_0a)\left[1-\frac{\rho_0}{\rho_1}\right]f_n'(a). \tag{47}$$

Substituting $f_n'(a)=-f_n(a)\tilde{t}_n/a$ in both these expressions, we obtain two linear equations which are then solved for $f_n(a)$ and $\tilde{I}_n(a)$. With the use of the Wronskian relation

$$j_n(k_0a)h_n'(k_0a)-j_n'(k_0a)h_n(k_0a)=-i/k_0^2a^2, \tag{48}$$

the end result is

$$f_n(a)=-af_n'(a)/\tilde{t}_n=\frac{(-i)^n(2n+1)}{ik_0a\left[k_0ah_n'(k_0a)+\frac{\rho_0}{\rho_1}h_n(k_0a)\tilde{t}_n\right]}, \tag{49}$$

$$\tilde{I}_n(a)=\frac{(-i)^n(2n+1)[k_0aj_n'(k_0a)+j_n(k_0a)\tilde{t}_n]}{ik_0a\left[k_0ah_n'(k_0a)+\frac{\rho_0}{\rho_1}h_n(k_0a)\tilde{t}_n\right]}. \tag{50}$$

Substituting these two results and the expression (41) of T_n in terms of $\tilde{I}_n(a)$ in (37), we get $\Delta_n=0$ or

$$(-i)^n(2n+1)+ik_0a^2h_n(k_0a)\left(\frac{\rho_0}{\rho_1}-1\right)f_n'(a)-ik_0^2a^4\sum_{q=1}^{\infty}\frac{C_{nq}G_{nq}}{\gamma_{nq}^2-k_0^2a^2}=0. \tag{51}$$

Therefore, with the choice (45) for $\tilde{t}_n$, the expressions (37) and (42) are significantly simplified as follows:

$$B_{n\mu}=\frac{a^2G_{n\mu}}{\gamma_{n\mu}^2-k_0^2a^2},\quad n=0,1,...,\quad \mu=1,2,..., \tag{52}$$

$$N_{n\mu}G_{n\mu}=\sum_{q=1}^{\infty}\frac{G_{nq}}{\gamma_{nq}^2-k_0^2a^2}[k_0^2a^2Q_2(n|q,\mu)-a\gamma_{nq}Q_4(n|q,\mu)],\quad n=0,1,...,\quad \mu=1,2,.... \tag{53}$$

Now, the expansion coefficients $G_{n\mu}$ are obtained by solving the homogeneous system of Eq. (53) together with the inhomogeneous equation (51). With $G_{n\mu}\sim 1/\mu$ as $\mu\to\infty$, in general we deduce from (52) that $B_{n\mu}\sim 1/\mu^3$. All these facts were verified by the numerical results obtained in the special cases considered in Sec. VII, i.e., for specific $b(r)$, $\rho(r)$.

It is worth remarking that in Ref. 1 the evaluation of $\tilde{t}_n$ from (45) was based on the value of $f_n'(a-)$ obtained from its relation to $f_n'(a+)$ via the boundary conditions at $r=a$; for the latter a series expression converging as $1/\mu^4$ was obtained for the exterior field. Although the same method could be followed here, too, the procedure adopted here provides an alternative method of evaluating the $B_{n\mu}$'s with optimum convergence. Both procedures were followed in Sec. VII for specific $b(r)$, $\rho(r)$ chosen to obtain numerical results. The results for the $B_{n\mu}$'s were identical.

One last remark should be added here: Since $f_n(a)$ and $f_n'(a)$ are unknown functions, the optimum value $\tilde{t}_n$ is not immediately available for starting the numerical solution on the basis of the reduced equations (51)–(53). Therefore, one must still start with an arbitrary value of t_n and via (42) and (40) obtain the optimum value $\tilde{t}_n$ from (45). As in Ref. 1, one may then repeat the solution with this value of $\tilde{t}_n$ on the basis of the reduced equations; the advantage gained is higher accuracy in the determination of $B_{n\mu}$.

VI. INHOMOGENEITIES DEPENDENT ON r, θ

Let us now assume that $b(\mathbf{r})=b(r,\theta)$ and $\rho(\mathbf{r})=\rho(r,\theta)$; the more general case of r-, θ-, φ dependence will be discussed briefly later; it is not considered here to help avoid long expressions and algebraic complexity. Continuity of $b(r,\theta)$, $\rho(r,\theta)$ at the center of the sphere requires that $b(0,\theta)$, $\rho(0,\theta)$ are constant, independent of θ. Also, the discontinuity of $\rho(r,\theta)$ at $r=a$ depends, in general, on θ: $\rho(a,\theta)=\rho_1(\theta)\neq\rho_0$, and we may set

$$\rho(r,\theta)=\rho_c(r,\theta)+[\rho_1(\theta)-\rho_0]u(a-r), \tag{54}$$

$$\frac{\partial\rho(r,\theta)}{\partial r}=\frac{\partial\rho_c(r,\theta)}{\partial r}+[\rho_0-\rho_1(\theta)]\delta(a-r). \tag{55}$$

These two further imply that

$$\rho_c(0,\theta)=\text{constant}-\rho_1(\theta) \quad \text{and} \quad \rho_c(a,\theta)=\rho_0. \tag{56}$$

Substituting now (3), (7), and (8) in (1), using the second expression of (4) in the integrand of (1), and the orthogonality relations for $P_n^m(\cos\theta)e^{im\varphi}$, we get for $n\geqslant|m|$, $m=0,\pm1,...$

$$\begin{aligned}f_{nm}(r)=d_{nm}(r)-\frac{ik_0}{2}(2n+1)\frac{(n-m)!}{(n+m)!}&\\ \times\Bigg\{\int_0^a\int_0^\pi j_n(k_0r_<)h_n(k_0r_>)P_n^m(\cos\theta')&\\ \times\sum_{p=|m|}^{\infty}g_{pm}(r')P_p^m(\cos\theta')r'^2\sin\theta'\,dr'\,d\theta'&\\ -a^2j_n(k_0r)h_n(k_0a)\int_0^\pi P_n^m(\cos\theta'),&\\ \times\frac{\rho_0-\rho_1(\theta')}{\rho(a,\theta')}\sum_{p=|m|}^{\infty}f'_{pm}(a)P_p^m(\cos\theta')\sin\theta'\,d\theta'\Bigg\},&\end{aligned} \tag{57}$$

where we have set

$$\begin{aligned}k_0^2\left[\frac{b(r,\theta)\rho(r,\theta)}{b_0\rho_0}-1\right]\sum_{p=|m|}^{\infty}f_{pm}(r)P_p^m(\cos\theta)&\\ -\frac{1}{\rho(r,\theta)}\frac{\partial\rho_c(r,\theta)}{\partial r}\sum_{p=|m|}^{\infty}f'_{pm}(r)P_p^m(\cos\theta)&\\ -\frac{1}{r^2\rho(r,\theta)}\frac{\partial\rho(r,\theta)}{\partial\theta}\sum_{p=|m|}^{\infty}f_{pm}(r)\frac{\partial P_n^m(\cos\theta)}{\partial\theta}&\\ =\sum_{p=|m|}^{\infty}g_{pm}(r)P_p^m(\cos\theta),&\end{aligned} \tag{58}$$

where the last term in (57) comes from the delta function $\delta(a-r')$ of (55) leading to a simple integration with respect to r', while $\rho(a,\theta')=\rho_1(\theta')$ and $d_n(r)=(-i)^n(2n+1)j_n(k_0r)$, as in (24).

We may remark at this point that the more general case in which ρ and b possess an additional φ dependence would require replacing (58) by (11), while the simple sum over p in (57) would be replaced by a double one. However, apart from the obvious additional algebraic complexity, the basic steps of the procedure remain the same.

Performing the integration in (57) yields

$$\begin{aligned}f_{nm}(r)=(-i)^n(2n+1)j_n(k_0r)-ik_0[h_n(k_0r)I_{nm}(r)&\\ +j_n(k_0r)L_{nm}(r)]-\frac{ik_0a^2}{2}(2n+1)\frac{(n-m)!}{(n+m)!}&\\ \times j_n(k_0r)h_n(k_0a)\sum_{p=|m|}^{\infty}\Gamma_{np}^m f'_{pm}(a),&\end{aligned} \tag{59}$$

with $I_{nm}(r)$, $L_{nm}(r)$ as in (10) and

$$\begin{aligned}\Gamma_{np}^m&=\int_0^\pi\frac{\rho_1(\theta)-\rho_0}{\rho_1(\theta)}P_n^m(\cos\theta)P_p^m(\cos\theta)\sin\theta\,d\theta\\ &=\frac{2}{2n+1}\frac{(n+m)!}{(n-m)!}\delta_{np}-\rho_0\\ &\quad\times\int_0^\pi\frac{1}{\rho_1(\theta)}P_n^m(\cos\theta)P_p^m(\cos\theta)\sin\theta\,d\theta.\end{aligned} \tag{60}$$

The constants Γ_{np}^m can be evaluated, in general, by numerical integration. If $\rho_1(\theta)=\rho_1=\text{constant}\neq\rho_0$, we get

$$\Gamma_{np}^m=\delta_{np}\frac{\rho_1-\rho_0}{\rho_1}\frac{2}{2n+1}\frac{(n+m)!}{(n-m)!}. \tag{61}$$

Differentiating (59) with respect to r, we further obtain

$$\begin{aligned}f'_{nm}(r)=(-i)^n(2n+1)k_0j'_n(k_0r)-ik_0^2[h'_n(k_0r)I_{nm}(r)&\\ +j'_n(k_0r)L_{nm}(r)]-\frac{ik_0^2a^2}{2}(2n+1)&\\ \times\frac{(n-m)!}{(n+m)!}j'_n(k_0r)h_n(k_0a)\sum_{p=|m|}^{\infty}\Gamma_{np}^m f'_{pm}(a).&\end{aligned} \tag{62}$$

As before, owing to the special form of $h_n(k_0r)I_{nm}(r)+j_n(k_0r)L_{nm}(r)$ in (10), no derivatives of the unknown functions $g_{nm}(r)$ enter into the expressions. Next, we expand the radial functions in Dini series following the steps of Sec. IV

$$\begin{aligned}g_{nm}(r)&=\sum_{\mu=1}^{\infty}G_{n\mu}^m j_n\left(\gamma_{n\mu}\frac{r}{a}\right),\\ f_{nm}(r)&=\sum_{\mu=1}^{\infty}B_{n\mu}^m j_n\left(\gamma_{n\mu}\frac{r}{a}\right).\end{aligned} \tag{63}$$

Substitution of (63) together with expressions equivalent to (29)–(31) in (59) and (62) and the orthogonality of $j_n[\gamma_{n\mu}(r/a)]$ in $0\leqslant r\leqslant a$ leads to equivalent to (37), (38) relations

$$B_{n\mu}^m=\Delta_n^m J_{n\mu}+\frac{a^2G_{n\mu}^m}{\gamma_{n\mu}^2-k_0^2a^2}, \tag{64}$$

$$\Delta_n^m=(-i)^n(2n+1)-\frac{ik_0a^2}{2}(2n+1)\frac{(n-m)!}{(n+m)!}$$

$$\times h_n(k_0a)\sum_{p=|m|}^{\infty}\Gamma_{np}^m f'_{pm}(a)-ik_0^2a^4T_n^m, \tag{65}$$

where T_n^m is the same as T_n in (38) or (41) if $G_{n\mu}^m$, $I_{nm}(a)$ replace $G_{n\mu}$, $\tilde{I}_n(a)$, respectively.

Also, without differentiating $f_{nm}(r)$ in (63), taking into account that $L_{nm}(a)=0$ and the expression (39) for $I_{nm}(a)$, we get from (62)

$$f'_{nm}(r)=k_0\Delta_n^m j'_n(k_0r)+a\sum_{\mu=1}^{\infty}\frac{G_{n\mu}^m\gamma_{n\mu}j'_n\left(\gamma_{n\mu}\frac{r}{a}\right)}{\gamma_{n\mu}^2-k_0^2a^2}. \tag{66}$$

Finally, we substitute the Dini expansions (63) for $g_{nm}(r)$, $f_{nm}(r)$ and expression (66) for $f'_{nm}(r)$ in (58); we then invoke the orthogonality of $j_n[\gamma_{n\mu}(r/a)]$ $(\mu=1,2,...)$ in $0\leqslant r\leqslant a$ to obtain

$$\frac{2N_{n\mu}}{2n+1}G_{n\mu}^m=\sum_{\mu'=1}^{\infty}\sum_{p=|m|}^{\infty}\left\{\Delta_p^m J_{p\mu'}+\frac{a^2G_{n\mu'}^m}{\gamma_{n\mu'}^2-k_0^2a^2}\right\}$$
$$\times Q_2^m(n,p|\mu,\mu')-k_0\sum_{p=|m|}^{\infty}\Delta_p^m Q_3^m(n,p|\mu)$$
$$-a\sum_{\mu'=1}^{\infty}\sum_{p=|m|}^{\infty}\frac{G_{p\mu'}^m\gamma_{p\mu'}}{\gamma_{p\mu'}^2-k_0^2a^2}Q_4^m(n,p|\mu,\mu'), \tag{67}$$

where, corresponding to (33)–(35), we have defined

$$Q_2^m(n,p|\mu,\mu')=\int_0^a\int_0^{\pi}j_n\left(\gamma_{n\mu}\frac{r}{a}\right)j_p\left(\gamma_{p\mu'}\frac{r}{a}\right)$$
$$\times\left\{k_0^2\left[\frac{b(r,\theta)\rho(r,\theta)}{b_0\rho_0}-1\right]P_p^m(\cos\theta)\right.$$
$$\left.-\frac{1}{r^2\rho(r,\theta)}\frac{\partial\rho(r,\theta)}{\partial\theta}\frac{\partial P_p^m(\cos\theta)}{\partial\theta}\right\}$$
$$\times P_n^m(\cos\theta)r^2\sin\theta\,d\theta\,dr, \tag{68}$$

$$Q_3^m(n,p|\mu)=\int_0^a\int_0^{\pi}j_n\left(\gamma_{n\mu}\frac{r}{a}\right)j'_p(k_0r)$$
$$\times\frac{1}{\rho(r,\theta)}\frac{\partial\rho_c(r,\theta)}{\partial r}P_p^m(\cos\theta)$$
$$\times P_n^m(\cos\theta)r^2\sin\theta\,d\theta\,dr, \tag{69}$$

$$Q_4^m(n,p|\mu,\mu')=\int_0^a\int_0^{\pi}j_n\left(\gamma_{n\mu}\frac{r}{a}\right)j'_p\left(\gamma_{n\mu'}\frac{r}{a}\right)$$
$$\times\frac{1}{\rho(r,\theta)}\frac{\partial\rho_c(r,\theta)}{\partial r}P_p^m(\cos\theta)$$
$$\times P_n^m(\cos\theta)r^2\sin\theta\,d\theta\,dr. \tag{70}$$

These coefficients Q_2^m, Q_3^m, Q_4^m can be evaluated, for each $m=0,\pm1,\pm2,...$, by numerical integration of the double integrals, while the $G_{n\mu}^m$ $(m=0,\pm1,\pm2,...,n\geqslant|m|,\mu=1,2,...)$ are obtained after truncation of the linear system of Eq. (67). What still remains to be found are the values of $f'_{pm}(a)$ and $\Sigma_{p=|m|}^{\infty}\Gamma_{np}^m f'_{pm}(a)$, which appear in Δ_n^m [see (65)] and depend on the $G_{n\mu}^m$'s. The simple case $\rho_1(\theta)=\rho_1=$constant, mentioned in (61), leads to

$$\sum_{p=|m|}^{\infty}\Gamma_{np}^m f'_{pm}(a)=\Gamma_{nn}^m f'_{nm}(a)$$
$$=\frac{\rho_1-\rho_0}{\rho_1}\frac{2}{2n+1}\frac{(n+m)!}{(n-m)!}. \tag{71}$$

Then, Δ_n^m obtains the form of Δ_n in (37); also (40) for $f'_{nm}(a)$ can be used and the solution proceeds much like the preceding θ-independent case. Even for each n, Δ_n^m can be made zero by properly selecting $t_{n,\text{optimum}}=[af'_{nm}(a)]/[f_{nm}(a)]\equiv\tilde{t}_n$; this leads to a simplified set of equations (67) and to improved convergence of all Dini series that appear and of the sequence $B_{n\mu}^m$ $(\mu=1,2,...)$.

When $\rho_1(\theta)$ is not constant the previous simplifications are no longer possible. The constants $f'_{nm}(a)$ are not so easy to eliminate from the system (67). Starting from (66), however, we obtain for $r=a$

$$f'_{nm}(a)=k_0j'_n(k_0a)\left\{(-i)^n(2n+1)-\frac{ik_0a^2}{2}(2n+1)\right.$$
$$\times\frac{(n-m)!}{(n+m)!}h_n(k_0a)\sum_{p=|m|}^{\infty}\Gamma_{np}^m f'_{pm}(a)$$
$$\left.-ik_0^2a^4\sum_{\mu=1}^{\infty}\frac{C_{n\mu}G_{n\mu}}{\gamma_{n\mu}^2-k_0^2a^2}\right\}$$
$$+a\sum_{\mu=1}^{\infty}\frac{G_{n\mu}^m\gamma_{n\mu}j'_n(\gamma_{n\mu})}{\gamma_{n\mu}^2-k_0^2a^2}. \tag{72}$$

For $n=|m|,|m|+1,...$ this relation provides a set of equations relating linearly the $f'_{nm}(a)$ $(n\geqslant|m|)$ to the $G_{n\mu}^m$. The simpler way to handle the problem is then to consider the $f'_{nm}(a)$ $(n\geqslant|m|)$ as additional unknowns, which, together with $G_{n\mu}^m$ $(\mu=1,2,...)$, form an enlarged set of unknown coefficients that can be determined from the enlarged set of equations (67) and (72). Apart from computer time considerations this enlarged system is no more difficult to handle numerically than the shorter system (67), that occurs when $\rho_1(\theta)=\rho_1=$constant.

Before closing this section, two remarks are in order. The first is related to the question of the boundary conditions at $r=a$. For $r\geqslant a$ we again may write

$$\tilde{\Phi}(\mathbf{r})=\sum_n\sum_m\tilde{f}_{nm}(r)P_n^m(\cos\theta)e^{im\varphi},\quad r\geqslant a, \tag{73}$$

and as in (18b) easily obtain from (1)

$$\tilde{f}_{nm}(r)=(-i)^n(2n+1)j_n(k_0r)-ik_0h_n(k_0r)I_{nm}(a)$$
$$-\frac{ik_0a^2}{2}(2n+1)\frac{(n-m)!}{(n+m)!}j_n(k_0a)$$
$$\times\sum_{p=|m|}^{\infty}\Gamma_{np}^m f'_{pm}(a)h_n(k_0r). \tag{74}$$

Differentiating this, we obtain $\tilde{f}'_{nm}(r)$, and setting $r=a$, $\tilde{f}_{nm}(a)$ and $\tilde{f}'_{nm}(a)$. The boundary conditions are[1]

$$\Phi(a,\theta)=\tilde{\Phi}(a,\theta)\ \text{ and }\ \frac{1}{\rho_1(\theta)}\frac{\partial\Phi(r,\theta)}{\partial r}\bigg|_{r=a}$$
$$=\frac{1}{\rho_0}\frac{\partial\tilde{\Phi}(r,\theta)}{\partial r}\bigg|_{r=a}. \tag{75}$$

The first implies $f_{nm}(a)=\tilde{f}_{nm}(a)$ and it is satisfied, if one uses (62) and (74) for $r=a$. The second requires that

$$\sum_{n=|m|}^{\infty}f'_{nm}(a)P_n^m(\cos\theta)\frac{1}{\rho_1(\theta)}=\frac{1}{\rho_0}\sum_{n=|m|}^{\infty}\tilde{f}'_{nm}(a)P_n^m(\cos\theta), \tag{76}$$

and it is satisfied if, besides $f_{nm}(a)=\tilde{f}_{nm}(a)$, one invokes (60), defining the constants Γ_{np}^m, and the Wroskian of the spherical Bessel functions.

This important test provides an irrefutable proof of the correctness of our approach when handling the delta function behavior of $\rho'(r)$ at $r=a$, namely, the incorporation of (54) and, mainly, (55) into the integral equation (1).

The second remark, related to this last test, leads to the conclusion that an exactly similar approach can be used to solve the integral equation (1) when the discontinuity of $\rho(r,\theta)$ occurs at an intermediate point $r=r_0$ $(0<r_0<a)$, instead of at $r=a$. Step discontinuities of $b(r,\theta)$ are of no consequence since no derivatives like $[\partial b(r)]/\partial r$ or $\nabla b(\mathbf{r})$ appear in (1).

VII. NUMERICAL RESULTS AND DISCUSSION

Numerical results were obtained and plotted for both radial and r-, θ dependence of the inhomogeneities. As in Ref. 1, for the simpler radial dependence we plot the magnitude of the total field $|\Phi|$ inside the inhomogeneity, versus r/a for three values of θ $(=0,\ \pi/2,$ and $\pi)$. Again, the maximum of $|\Phi|$ is close to the maximum appearing in the curves of $\theta=0,\ \pi$, while its minimum is close to that corresponding to the curve for $\theta=\pi/2$. This may be related to the increased directionality of the radiation pattern of the external field at such wave numbers.[10] So, the above plots provide a good idea of field penetration in the sphere.

At first in Figs. 1 and 2 we plot $|\Phi(0,0,z)|$ $(\theta=0,\pi)$ and $|\Phi(x,y,0)|$ $(\theta=\pi/2)$, respectively, for various sizes k_0a in the case of so-called Gaussian spheres [Ref. 11, p. 2018]. In this case the special radial dependence of $\rho(r)$ and $k(r)$ allows an exact analytical solution, based on the separation of variables; so it may serve as a test of our more general approach. We have chosen, in particular, $\rho(r)=20\rho_0\exp(-200r^2)$, $k^2(r)=1000+40\,000r^2$, $0\leqslant r\leqslant 0.1=a$. Our results are identical with the analytical ones obtained via Martin's analysis,[11] confirming the correctness of our procedure. The agreement is so good that it makes the corresponding plots in Figs. 1, 2 indistinguishable from one another.

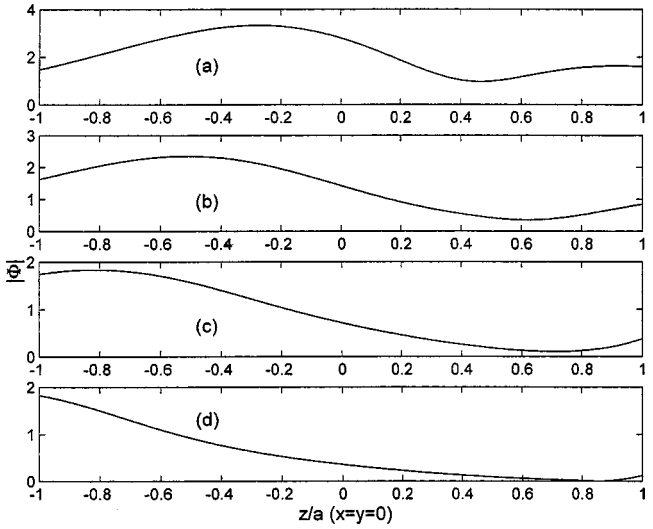


FIG. 1. Total field $|\Phi(0,0,z)|$ inside a spherical inhomogeneity described by $\rho(r)=20\rho_0\exp(-200r^2)$, $k^2(r)=1000+40\,000r^2$, $0\leqslant r\leqslant 0.1=a$, with $k_0a=$(a) 2.0958; (b) 4.191 69; (c) 8.3834; (d) 16.7668.

In Figs. 3 and 4 we next present the same results for $\rho(r)=\rho_0[3+2\cos(\pi(r/a))]$, $b(r)=b_0[7.5-2.5\cos(\pi(r/a))]$, $0\leqslant r/a\leqslant 1$. Here, contrary to the preceding case, the density $\rho(r)$ is continuous at $r=a$. In all these cases we have let the size of k_0a vary from about 2 to about 17.

From the results it seems that, at least in the particular cases studied here, when the inhomogeneity of the sphere is not varying sharply the field in it exhibits relatively small peak values. These become smaller the greater the difference of the characteristic parameter ka of this region from that of the outer space. On the other hand, in cases of sharp inhomogeneities the induced field presents sharp peak values,

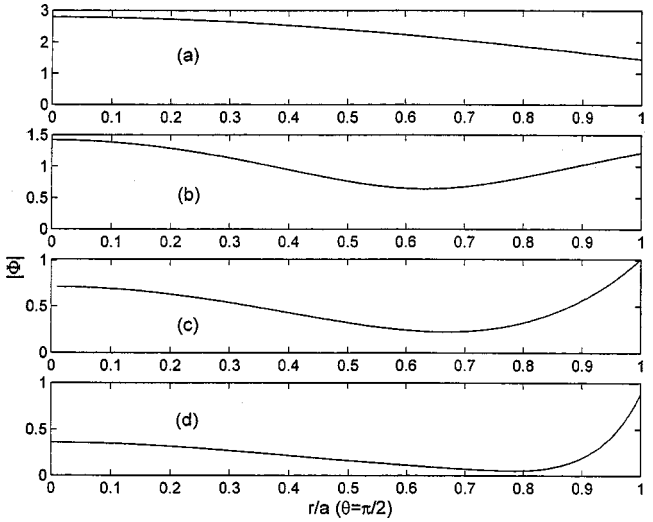


FIG. 2. Total field $|\Phi(x,y,0)|$ inside a spherical inhomogeneity described by $\rho(r)=20\rho_0\exp(-200r^2)$, $k^2(r)=1000+40\,000r^2$, $0\leqslant r\leqslant 0.1=a$, with $k_0a=$(a) 2.0958; (b) 4.191 69; (c) 8.3834; (d) 16.7668.

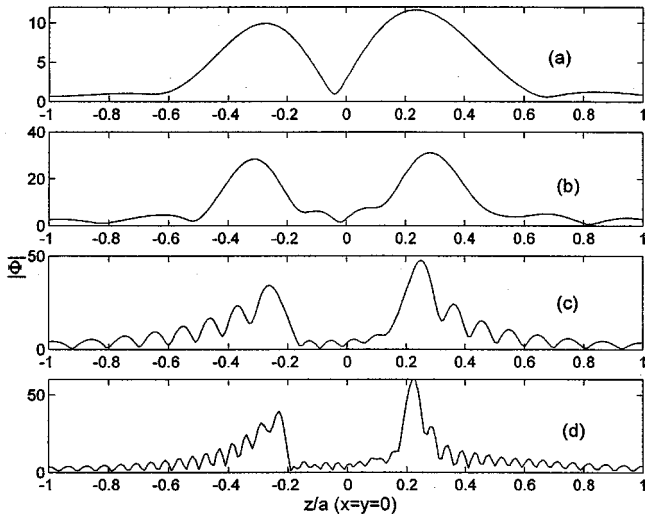


FIG. 3. Total field $|\Phi(0,0,z)|$ inside a spherical inhomogeneity described by $\rho(r)=\rho_0[3+2\cos(\pi(r/a))]$, $b(r)=b_0[7.5-2.5\cos(\pi(r/a))]$, $0\leqslant r/a\leqslant 1$, with $k_0a=$(a) 2.0958; (b) 4.191 69; (c) 8.3834; (d) 16.7668.

becoming higher for more rapid variations. All these may be attributed to localization phenomena and to trapping of the incident energy.

The rate of convergence of the Dini expansions and the corresponding matrix sizes are very close to those in Ref. 1, for the same k_0a, and will not be repeated here. Suffice to say that for $k_0a=16.7668$ truncation of the index n to 25 and of μ to 36 (smaller for higher n) provide for $|\Phi|$ accuracy of approximately 0.1%. For $k_0a=2.0958$ the corresponding values are $n_{\max}=7$ and $\mu_{\max}=7$ (smaller for higher n). Again, in these cases of simple radial dependence the method of choosing the optimum value of t_n, which improves the convergence and the accuracy of the procedure, follows the steps developed in Ref. 1. As explained further in connection with (44) and (45), the values of $f_n(a)$ and $f_n'(a)$ can be obtained via (46) and (47) without recourse to the boundary conditions.

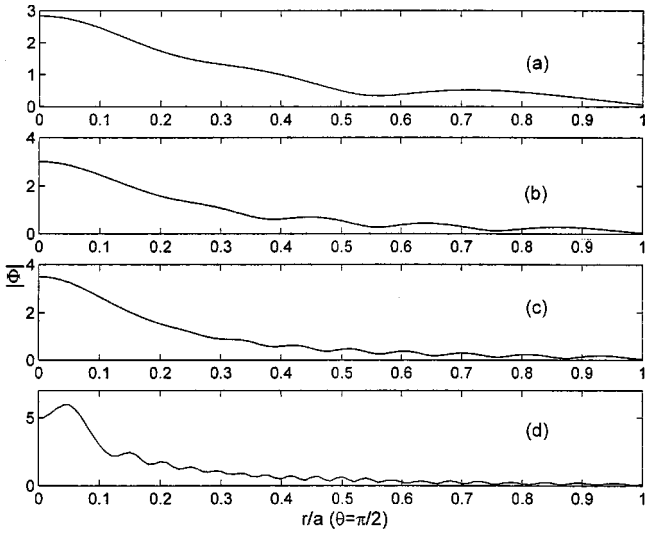


FIG. 4. Total field $|\Phi(x,y,0)|$ inside a spherical inhomogeneity described by $\rho(r)=\rho_0[3+2\cos(\pi(r/a))]$, $b(r)=b_0[7.5-2.5\cos(\pi(r/a))]$, $0\leqslant r/a\leqslant 1$, with $k_0a=$(a) 2.0958; (b) 4.191 69; (c) 8.3834; (d) 16.7668.

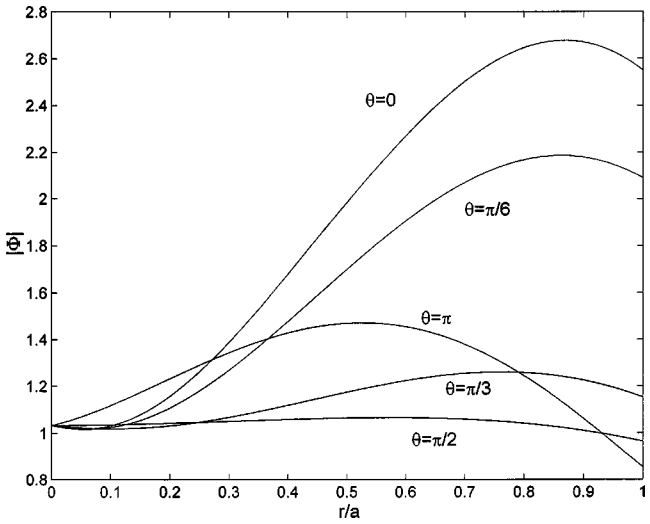


FIG. 5. Total field $|\Phi|$ inside a spherical inhomogeneity described by $\rho(r)=\rho_0\exp[(r/a)((r/a)-1)\cos\theta]$, $b(\mathbf{r})\rho(\mathbf{r})=1+[0.65-0.25\cos(\pi(r/a))](r/a)\times(2-\sin\theta)$, with $k_0a=2.0958$.

Finally, in Fig. 5, $|\Phi|$ is plotted for certain values of θ in the case

$$\rho(r)=\rho_0\exp\left[\frac{r}{a}\left(\frac{r}{a}-1\right)\cos\theta\right]$$

$$\frac{b(r)\rho(r)}{b_0\rho_0}=1+\left[0.65-0.25\cos\left(\pi\frac{r}{a}\right)\right]\frac{r}{a}(2-\sin\theta)$$

$$0\leqslant\frac{r}{a}\leqslant 1 \tag{77}$$

Here, the small value of $k_0a=2.0958$ was chosen to reduce matrix size, which with θ dependence of ρ, b increases significantly compared with the simple radial dependence. Further results, for other values of θ, show that they fall in between the plotted curves. So, those are sufficient for a pretty accurate indication of field variation inside the sphere.

VIII. CONCLUSIONS

Acoustic fields induced by external sources in media with both inhomogeneous density $\rho(\mathbf{r})$ and compressibility $b(\mathbf{r})$ have been treated by the direct hybrid (analytical–numerical) method initially developed in Ref. 1 for media with constant density ρ_0. This important extension is characterized by two serious complications, absent in Ref. 1. The volume integral equation contains now derivatives of $\rho(\mathbf{r})$, $\Phi(\mathbf{r})$ which, unless properly handled, may appear also in the unknown functions $g_{nm}(r)$ of the integrand, Eq. (6). First, by the simple vector identity (4) we avoid second derivatives of $\Phi(\mathbf{r})$, expressing the integrand in terms of $\Phi(\mathbf{r})$ and $\nabla\Phi(\mathbf{r})$. Second, based on the expressions (9), (10) (following the analysis in Ref. 6) we develop for both the radial expansion functions $f_{nm}(r)$ of $\Phi(\mathbf{r})$ and for $f_{nm}'(r)$, respectively, the expansions (59) and (62) in terms of the integrand-expansion-functions $g_{nm}(r)$ only, not their derivatives. This is a clear advantage of our analytical approach over other possible numerical methods that cannot avoid derivatives of $\Phi(\mathbf{r})$ [and their complicated relations with values of $\Phi(\mathbf{r})$] in the treatment of the volume integral. This successful treat-

ment is further implemented by introducing explicitly in (54) and (55) the step- and the delta-function discontinuities of $\rho(\mathbf{r})$, $[\partial\rho(\mathbf{r})]/\partial$r, respectively, at $r=a$. We believe that this constitutes a significant approach. The whole procedure is tested analytically against the general boundary conditions at $r=a$, for any $\rho(r,\theta)$, $b(r,\theta)$; this successful test establishes beyond any doubt the validity of our approach. Furthermore, the same procedure may be used to treat either discontinuities of $\rho(\mathbf{r})$ inside the sphere or the more difficult electromagnetic problem. Again, the restriction to spherical shapes, as in Ref. 1, allows use of Dini's expansions of a general type for the radial functions, with superior convergence over other possible sets of orthogonal expansions as well as the possibility of further optimization of the convergence of the Dini expansions.

ACKNOWLEDGMENT

The present work has been supported by the program of basic research ''Archimides'' of the Institute of Communication and Computer Systems, National Technical University of Athens, Athens 15773, Greece, as well as by the Embirikion Foundation of Greece.

[1] G. C. Kokkorakis, J. G. Fikioris, and G. Fikioris, ''Field induced in inhomogeneous spheres by external sources. I. The scalar case,'' J. Acoust. Soc. Am. **112**(4), 1297–1306 (2002).
[2] W. C. Chew, *Waves and Fields in Inhomogeneous Media* (Van Nostrand Reinhold, New York, 1990).
[3] M. A. Jensen, ''A recursive Green's function technique for acoustic scattering from heterogeneous objects,'' J. Acoust. Soc. Am. **103**(2), 713–720 (1998).
[4] A. N. Ivakin, ''A unified approach to volume and roughness scattering,'' J. Acoust. Soc. Am. **103**(2), 827–837 (1998).
[5] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
[6] J. G. Fikioris, ''Electromagnetic field of in the source region of continously varying current density,'' Q. Appl. Math. **LIV**, No. 2, 201–209 (1996).
[7] C. Müller, *Foundations of the Mathematical Theory of Electromagnetic Waves* (Springer, Berlin, 1969).
[8] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed. (Cambridge University Press, Cambridge, England, 1958).
[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
[10] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
[11] P. A. Martin, ''Acoustic scattering by inhomogeneous spheres,'' J. Acoust. Soc. Am. **111**(5), 2013–2018 (2002).

Scattering of sound from axisymetric sources by multiple circular cylinders

Scott E. Sherer[a)]
Air Vehicles Directorate, Air Force Research Laboratory, Wright-Patterson AFB, Ohio 45433-7521

(Received 10 January 2003; revised 15 November 2003; accepted 24 November 2003)

A general analytic method for calculating the scattering of sound by multiple rigid circular cylinders arranged in an arbitrary parallel configuration is presented. The sound scattered by this collection of cylinders is generated by a time-periodic, spatially distributed, axisymmetric source located within the domain of interest. A Hankel transform method is used to calculate the incident field, while separation of variables is used to obtain the scattered fields from each cylinder in the collection. The unknown scattering coefficients are determined through the use of general addition theorems that allows the various fields to be readily transformed between coordinate systems. The method is validated using various two-, three-, and four-cylinder configurations, and the number of coefficients that must be retained in the truncated series is examined. Benchmark configurations consisting of two- and three-cylinder systems with cylinders of varying radii are also presented. These solutions have been used to validate computational aeroacoustic solvers developed for complex geometries. [DOI: 10.1121/1.1641790]

PACS numbers: 43.20.Fn [GCG] Pages: 488–496

I. INTRODUCTION

The scattering of waves by two-dimensional circular cylinders is a fundamental problem in both acoustics and electromagnetics. Analytical solutions to the scattering of plane waves by a single circular cylinder are presented by, among others, Morse and Ingard[1] within the context of acoustics and Balanis[2] in electromagnetics. Single-cylinder solutions have also been developed for incident fields other than plane waves. For example, Shenderov[3] presents a solution for the scattering of sound generated by a cylindrical line source from a single, rigid circular cylinder, while Balanis[2] presents the solution for the equivalent electromagnetic problem. Various authors have also developed solutions for single-cylinder scattering with spatially distributed acoustic sources. These include Kurbatskii,[4] who used a Green's function approach to calculate the scattering from a single, rigid, circular cylinder using a source with a Gaussian spatial distribution, and Morris,[5] who developed via a Hankel transform approach solutions for scattering by a single, acoustically permeable, circular cylinder for various spatially distributed sources including Gaussian and circular disk distributions.

Exact and approximate solutions have also been developed for the scattering of plane waves by multiple parallel circular cylinders. One of the earliest solutions to this problem for rigid cylinders was developed by Twersky,[6] who decomposed the total field into an incident field plus higher order scattered field contributions whose coefficients were determined in an iterative manner. Additional solution techniques were later developed by Young and Bertrand[7] as well as Peterson and Ström.[8] In the former, acoustic scattering calculations for a plane wave incident upon two identical, rigid, circular cylinders were performed using both direct matrix inversion and iterative procedures and compared to experimental data. The latter technique used a recursive approach based on the ''T-matrix method''[9] to formulate the scattering from an arbitrary number of general scatterers with a general, nonsingular incident field, although results are only presented for plane-wave scattering from two spheres. A general classification system for the various techniques available to solve the multiple-scattering problem as well as additional references for the multiple-cylinder case is given by Elsherbeni.[10]

Multiple-scattering phenomena have also received considerable attention as related to phononic crystals and wave localization.[11] Within this application, solution methods have been developed by Kafesaki and Economou,[12] which was used by Kafesaki *et al.*[13] to examine the phononic properties of air bubbles in water, and by Psarobas *et al.*,[14] which was employed by Yang *et al.*[15] in their investigation of the phononic properties of tungsten carbide beads in water. Similarly, exact solutions for the localization lengths associated with a random array of air cylinders in a water medium are presented by Gupta and Ye.[16]

In the current work, an analytic method is developed to obtain solutions to the problem of the scattering of sound generated by two types of axisymmetric sources by multiple rigid, circular cylinders of varying radii. The long axes of the cylinders are assumed to be parallel to each other, and thus the problem is two-dimensional. However, the distribution of the cylinders within this two-dimensional plane is arbitrary under the constraint that they do not overlap nor are they in contact with each other. This approach, which is an extension of the single-cylinder work of Morris,[5] utilizes a Hankel transform method to determine the incident field from a spatially distributed acoustic source. Separation of variables provides the general solution for the scattered field from each cylinder in its own coordinate system, and cylindrical addition theorems are employed to allow the application of the

a)Electronic mail: Scott.Sherer@wpafb.af.mil

boundary conditions in the various coordinate systems. The scattering coefficients for all cylinders are solved simultaneously by direct inversion, and thus the approach would be considered a boundary value solution method per the classification scheme of Ref. 10.

The solutions presented here are analytic in nature, although the necessary truncation of infinite series, numerical inversion of potentially ill-conditioned matrices, and the numerical evaluation of certain integrals can result in the generation of numerical error if not addressed. Because it is analytic, this approach is useful for establishing benchmark scattering solutions that may be used to assess the accuracy of computational aeroacoustic solvers developed for geometrically complex configurations.[17–19] The use of solutions generated by a spatially distributed source is preferred over a line source or plane-wave solution for this purpose due to the presence of a singularity in the total field for the former and the requirement to propagate the incident wave into the domain of interest from beyond its boundaries when using the total field formulation for the latter. This analytic approach may also prove amenable to parametric studies of the scattering behavior of systems involving multiple-cylinder geometries.

The following section addresses theoretical analysis, including the formulation of the problem under consideration, its analytic solution, and the acoustic source types considered. The next section contains results obtained for various configurations and source frequencies, including validation studies and investigations of some of the numerical issues associated with this procedure. This section concludes with scattered and total field results obtained for two- and three-cylinder benchmark configurations.

II. THEORETICAL ANALYSIS

A. Problem formulation

The problem considered here is the scattering of sound generated from a time-periodic, axisymmetric source by a set of M parallel, infinitely long (i.e., two-dimensional), rigid, circular cylinders. The arrangement of the cylinders and the source within the unbounded, two-dimensional domain of interest is arbitrary, with the caveat that the cylinders may not be in contact with nor overlap one another. This domain initially consists of a fluid with a speed of sound given by c_∞. A reference length l_{ref}, given by the diameter of the largest cylinder in the collection, and a reference time t_{ref}, given by $t_{\text{ref}}=l_{\text{ref}}/c_\infty$, have been used throughout this work for nondimensionalization purposes. It is also assumed that there is no mean flow in the domain.

The geometric relationships between any two cylinders i and j in this collection, the acoustic source centered at S, and an arbitrary observation point at O are given in Fig. 1. A polar coordinate system (r_i,θ_i) is defined based on the center of cylinder i, while a second polar coordinate system (r_j,θ_j) is defined based on the center of cylinder j. A third polar coordinate system (R,ϕ) is centered on the source. The positions of the cylinders with respect to the source are given in this source-based polar coordinate system as (L_i,α_i) and (L_j,α_j), respectively. The position of the cylinder j in the coordinate system of cylinder i is given by (D_{ij},ψ_{ij}), and likewise the position of the cylinder i in the coordinate system of cylinder j is given by (D_{ji},ψ_{ji}). It is noted that $D_{ij}=D_{ji}$ and $|\psi_{ij}-\psi_{ji}|=\pi$. The radii of the cylinders are denoted by a_i and a_j, respectively, and in general are not equal.

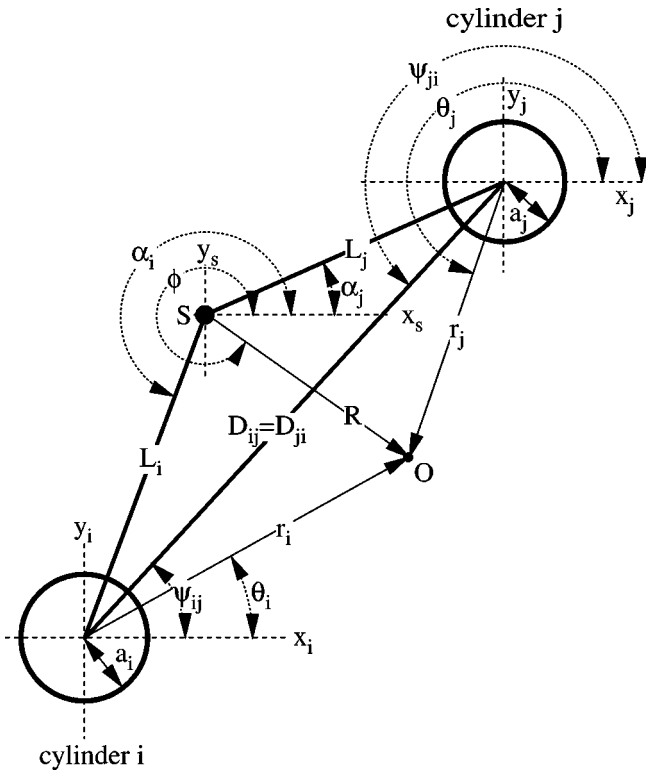


FIG. 1. Schematic showing coordinate system definitions and geometric relationships between acoustic source located at point S, an observation point at point O, and two arbitrary cylinders labeled i and j.

Assuming linear scattering and a time-dependent, axisymmetric source of the form $S(R,t)=f(R)e^{-\iota\omega t}$, the incident field satisfies the nonhomogeneous Helmholtz equation, here written in terms of the source coordinate system as

$$\frac{d^2\Phi^{\text{inc}}}{dR^2}+\frac{1}{R}\frac{d\Phi^{\text{inc}}}{dR}+k^2\Phi^{\text{inc}}=\tilde{f}(R), \tag{1}$$

where $\tilde{f}(R)=-\iota kf(R)$ and $k=\omega/c_\infty$. Note that the ϕ dependence has been dropped due to axisymmetry of the source. Similarly, the scattered field generated by each cylinder independently satisfies the homogeneous Helmholtz equation, written in the local, cylinder-based coordinate system as

$$\frac{\partial^2\Phi_i^s}{\partial r_i^2}+\frac{1}{r_i}\frac{\partial\Phi_i^s}{\partial r_i}+\frac{1}{r_i^2}\frac{\partial^2\Phi_i^s}{\partial\theta_i^2}+k^2\Phi_i^s=0,\quad i=1,2,...,M. \tag{2}$$

The principle of superposition is used to write the total field as the sum of the incident and scattered fields

$$\Phi=\Phi^{\text{inc}}+\sum_{i=1}^{M}\Phi_i^s. \tag{3}$$

The standard rigid-body boundary conditions applied at the surface of each cylinder j thus may be expressed in terms of these fields as

$$\sum_{i=1}^{M} \left.\frac{\partial \Phi_i^s}{\partial r_j}\right|_{r_j=a_j} = -\left.\frac{\partial \Phi^{\text{inc}}}{\partial r_j}\right|_{r_j=a_j} \qquad j=1,2,...,M. \tag{4}$$

B. Solution development

Morris[5] solved Eq. (1) using a Hankel transform method to determine the incident field, given in terms of the source-based coordinate system as

$$\Phi^{\text{inc}}(R) = -\int_0^\infty \frac{sJ_0(sR)\tilde{F}(s)}{(s^2-k^2)}ds, \tag{5}$$

where $\tilde{F}(s)$ is the Hankel transform of the spatial source distribution $\tilde{f}(R)$

$$\tilde{F}(s) = \int_0^\infty RJ_0(sR)\tilde{f}(R)dR, \tag{6}$$

and J_n is the Bessel function of the first kind of order n. Likewise, the solution for each scattered field may be obtained by solving Eq. (2) using separation of variables. Invoking an outgoing radiation boundary condition at infinity for the $e^{-\imath\omega t}$ time dependency, the scattered field from the ith cylinder is written in its local coordinate system as

$$\Phi_i^s(r_i,\theta_i) = A_{0i}H_0(kr_i) + \sum_{n=1}^{\infty} H_n(kr_i)[A_{ni}\cos(n\theta_i) + B_{ni}\sin(n\theta_i)] \tag{7}$$

where H_n is the Hankel function of the first kind of order n.

The scattered field coefficients A_{ni} and B_{ni} from Eq. (7) are found though the application of the M boundary conditions given by Eq. (4). This is accomplished through the application of the Graf addition theorem,[20] which allows the incident and scattered fields to be written in terms of the coordinate system of each cylinder in order to apply the boundary conditions at its surface. Using this theorem, the Bessel and Hankel functions in Eqs. (5) and (7) are written in terms of the coordinate system of cylinder j as

$$J_0(kR) = \sum_{m=0}^{\infty} \epsilon_m(-1)^m J_m(kL_j)J_m(kr_j) \times[\cos(m\alpha_j)\cos(m\theta_j) + \sin(m\alpha_j)\sin(m\theta_j)], \tag{8}$$

and

$$H_n(kr_i)\cos(n\theta_i) = \sum_{m=0}^{\infty} \frac{\epsilon_m}{2}\{[(-1)^m\mathcal{K}_{ijmn}^{c+} + \mathcal{K}_{ijmn}^{c-}]\cos(m\theta_j) + [(-1)^m\mathcal{K}_{ijmn}^{s+} - \mathcal{K}_{ijmn}^{s-}]\sin(m\theta_j)\}, \tag{9}$$

$$H_n(kr_i)\sin(n\theta_i) = \sum_{m=0}^{\infty} \frac{\epsilon_m}{2}\{[(-1)^m\mathcal{K}_{ijmn}^{s+} + \mathcal{K}_{ijmn}^{s-}]\cos(m\theta_j) - [(-1)^m\mathcal{K}_{ijmn}^{c+} - \mathcal{K}_{ijmn}^{c-}]\sin(m\theta_j)\}, \tag{10}$$

where ϵ_m is the Neumann factor, defined as

$$\epsilon_m = \begin{cases} 1 & m=0 \\ 2 & m\geqslant 1 \end{cases}, \tag{11}$$

and

$$\mathcal{K}_{ijmn}^{(c,s)\pm} = H_{n\pm m}(kD_{ij})J_m(kr_j)\cdot\begin{cases}\cos[(n\pm m)\psi_{ij}] \\ \sin[(n\pm m)\psi_{ij}]\end{cases}. \tag{12}$$

It is noted that the transformations for the Hankel functions are only valid for $r_j<D_{ij}$, but that this condition will always be satisfied on the surface of the cylinders where the boundary conditions are evaluated provided they are neither overlapping nor in contact with one another.

Using the previous transformations in conjunction with the incident and scattered fields given by Eqs. (5) and (7), the coefficients A_{ni} and B_{ni} may be determined through the application of the boundary conditions as given by Eq. (4) on the surface of each cylinder. The resulting M equations consist of infinite summations on both the left- and right-hand side of the equality, with each term possessing a factor of $\cos(m\theta_j)$, or of $\sin(m\theta_j)$, or having no θ_j dependency at all (the $m=0$ case). Equating on a term-by-term basis and truncating the infinite series at some finite value N yields $(2N+1)$ equations for each cylinder, or $M\cdot(2N+1)$ equations for the entire cylinder set. These equations are linear with respect to the $M\cdot(2N+1)$ unknown coefficients A_{ni} and B_{ni}, and may be expressed concisely in matrix form as

$$Z\mathbf{x} = \mathbf{b}. \tag{13}$$

The matrix Z may be written in terms of M^2 submatrices of dimension $(2N+1)\times(2N+1)$ as

$$Z = \begin{bmatrix} \mathcal{Z}_{11} & \mathcal{Z}_{12} & \cdots & \mathcal{Z}_{1M} \\ \mathcal{Z}_{21} & \mathcal{Z}_{22} & \cdots & \mathcal{Z}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{Z}_{M1} & \mathcal{Z}_{M2} & \cdots & \mathcal{Z}_{MM} \end{bmatrix}. \tag{14}$$

The off-diagonal submatrices $\mathcal{Z}_{ij}$, $i\neq j$, may be further subdivided as

$$\mathcal{Z}_{ij} = \begin{bmatrix} \mathcal{C}_{ij}^{+} & \mathcal{S}_{ij}^{+} \\ \mathcal{S}_{ij}^{-} & -\mathcal{C}_{ij}^{-} \end{bmatrix}, \tag{15}$$

where the elements of these submatrices are given by

$$[(\mathcal{C},\mathcal{S})_{ij}^{\pm}]_{mn} = \frac{\epsilon_m}{2}\overline{\mathcal{K}}_{jimn}^{(c,s)\pm} \rightarrow \begin{cases} \mathcal{C}^+:\{m=0\rightarrow N, n=0\rightarrow N\} \\ \mathcal{S}^+:\{m=0\rightarrow N, n=1\rightarrow N\} \\ \mathcal{S}^-:\{m=1\rightarrow N, n=0\rightarrow N\}, \\ \mathcal{C}^-:\{m=1\rightarrow N, n=1\rightarrow N\} \end{cases} \tag{16}$$

with

$$\overline{\mathcal{K}}_{ijmn}^{(c,s)\pm} = (-1)^m\mathcal{K}_{ijmn}^{(c,s)+\prime} \pm \mathcal{K}_{ijmn}^{(c,s)-\prime}, \tag{17}$$

and from Eq. (12)

$$\mathcal{K}_{ijmn}^{(c,s)\pm\prime}=\left.\frac{\partial \mathcal{K}_{ijmn}^{(c,s)\pm}}{\partial r_j}\right|_{r_j=a_j}$$
$$=\frac{k}{2}H_{m\pm n}(kD_{ij})[J_{m-1}(ka_j)-J_{m+1}(ka_j)]$$
$$\cdot\begin{Bmatrix}\cos[(n\pm m)\psi_{ij}]\\ \sin[(n\pm m)\psi_{ij}]\end{Bmatrix}. \tag{18}$$

The submatrices $\mathcal{Z}_{jj}$ on the diagonal of Eq. (14) are themselves diagonal matrices of the form

$$\mathcal{Z}_{jj}=\begin{bmatrix} H'_{0j} & & & & & & \\ & \ddots & & & & 0 & \\ & & H'_{Nj} & & & & \\ & & & H'_{1j} & & & \\ & 0 & & & \ddots & & \\ & & & & & & H'_{Nj}\end{bmatrix}, \tag{19}$$

where

$$H'_{nj}=\left.\frac{\partial H_n(kr_j)}{\partial r_j}\right|_{r_j=a_j}=\frac{k}{2}[H_{n-1}(ka_j)-H_{n+1}(ka_j)]. \tag{20}$$

The solution and right-hand-side vectors may also be expressed in block form as

$$\mathbf{x}=\begin{bmatrix}\boldsymbol{\mathcal{X}}_1\\ \boldsymbol{\mathcal{X}}_2\\ \vdots\\ \boldsymbol{\mathcal{X}}_M\end{bmatrix}\rightarrow\boldsymbol{\mathcal{X}}_i=\begin{bmatrix}A_{0i}\\ A_{1i}\\ \vdots\\ A_{Ni}\\ B_{1i}\\ B_{2i}\\ \vdots\\ B_{Ni}\end{bmatrix}, \tag{21}$$

and

$$\mathbf{b}=\begin{bmatrix}\boldsymbol{\mathcal{B}}_1\\ \boldsymbol{\mathcal{B}}_2\\ \vdots\\ \boldsymbol{\mathcal{B}}_M\end{bmatrix}\rightarrow\boldsymbol{\mathcal{B}}_j=\begin{bmatrix}I_{0j}^{c\prime}\\ I_{1j}^{c\prime}\\ \vdots\\ I_{Nj}^{c\prime}\\ I_{1j}^{s\prime}\\ I_{2j}^{s\prime}\\ \vdots\\ I_{Nj}^{s\prime}\end{bmatrix}, \tag{22}$$

with

$$I_{mj}^{(c,s)\prime}=\frac{\epsilon_m}{2}(-1)^m\int_0^\infty\frac{s^2J_m(sL_j)[J_{m-1}(sa_j)-J_{m+1}(sa_j)]\tilde{F}(s)}{s^2-k^2}ds\cdot\begin{Bmatrix}\cos(m\alpha_j)\\ \sin(m\alpha_j)\end{Bmatrix} \tag{23}$$

Upon solving this system of equations for the coefficients A_{ni} and B_{ni}, the scattered and total fields may be found via Eqs. (7) and (3), respectively.

C. Acoustic source types

To determine the incident field given by Eq. (5) and the right-hand-side vector **b** using the definitions in Eqs. (22) and (23), a specific form of $\tilde{F}(s)$ is required. Two particular source types will be considered here: a cylindrical line source and a spatially distributed cylindrical source with a Gaussian spatial distribution. Morris[5] previously considered both of these source types for the single-cylinder case, and the reader is referred there for more detail.

For the cylindrical line source, the spatial source function $f(R)$ and its corresponding Hankel-transformed counterpart $\tilde{F}(s)$ are given by

$$f(R)=\frac{\delta R}{2\pi R}\rightarrow\tilde{F}(s)=\frac{-\iota k}{2\pi}, \tag{24}$$

where $\delta(R)$ is the Dirac delta function. Integrating in the complex plane and applying the residue theorem yields for the incident field

$$\Phi^{\text{inc}}(R)=-\frac{k}{4}H_0(kR). \tag{25}$$

Likewise, the components making up the vector **b** are given by

$$I_{mj}^{(c,s)\prime}=\frac{\epsilon_m k^2}{8}(-1)^mH_m(kL_j)[J_{m-1}(ka_j)$$
$$-J_{m+1}(ka_j)]\cdot\begin{Bmatrix}\cos(m\alpha_j)\\ \sin(m\alpha_j)\end{Bmatrix}. \tag{26}$$

The second source type considered here is a spatially distributed source of the form

$$f(R)=\exp(-dR^2)\rightarrow\tilde{F}(s)=\frac{-\iota k}{2d}\exp\left(-\frac{s^2}{4d}\right). \tag{27}$$

For this source, the integrals in Eqs. (5) and (23) cannot be integrated analytically. The approach used by Morris[5] to perform the integration numerically is instead used with some minor modifications. In this approach, the variable of integration s is replaced by a new variable τ such that

$$s=\tau-\iota\gamma\exp[-\beta(\tau-k)^2], \tag{28}$$

where $\gamma=k/10$ and $\beta=-\ln(1\times10^{-10}/\gamma)/k^2$. The curve defined by the new variable τ is deformed below the real axis,

thus by-passing the pole located at $s=k$ and allowing the numerical integration to proceed. While in Ref. 5 a ten-point Gauss–Legendre integration scheme was employed over individual intervals of length $k/40$ along the integration path, here the numerical integration is performed by the software package MATHEMATICA.[21] This approach allows for intervals of various lengths to be employed while still maintaining a specified error tolerance. The use of larger intervals in those regions that can be numerically integrated by MATHEMATICA to the specified precision decreases the time it takes to calculate the value of the integral over the entire path. Even so, with $M\cdot(N+1)$ integrals required to construct the right-hand side of Eq. (13), the numerical integration of Eq. (23) is still the most time-consuming aspect of calculating the scattering coefficients. Note that the selection of this contour is not unique, and that other contours may be used to perform the numerical integration.

III. RESULTS AND DISCUSSION

All results presented in this work are normalized by the value of the incident field at a distance $R=L_1$ from the source, i.e., the distance from the source to the center of the left-most cylinder in a given configuration.

A. Implementation and validation

To obtain the incident and scattered fields for the various situations examined here, the previously presented formulation was coded in MATHEMATICA. This software package was used to solve the linear system of equations given by Eq. (13) and to perform the necessary numerical integrations for the cases involving the spatially Gaussian source. As discussed in the Introduction, there are several issues associated with the numerical implementation of this approach. These include the truncation of the infinite summations in the series solution for the scattered fields and in application of the addition theorems, the potential for the coefficient system to be ill-conditioned, and the numerical evaluation of the integrals required for the spatially Gaussian source. These issues will now be discussed.

The issue of ill-conditioning arises from the potential disparity in the eigenvalues of the coefficient matrix $\mathcal{Z}$ as additional terms are retained in the series solution for the scattered fields. Two methods were employed to address this issue. First, only the minimum number of scattering coefficients required to obtain a converged solution was retained, as the condition number of the coefficient matrix grows rapidly as additional terms are included. Second, partial pivoting was used in conjunction with Gaussian elimination to find the scattered coefficients, which helps to limit any round-off error.[22] To test the effectiveness of these steps, several of the cases shown in this work were solved using various levels of precision by first calculating the elements of the coefficient matrix and right-hand-side vector using exact algebra, and then evaluating them to a specified precision before solving the linear system. For all cases examined, the precision of the solutions obtained was comparable to the original precision used to generate the matrix elements, i.e., solutions generated using four digits of precision matched solutions obtained with higher precision to the first four decimal places. Thus, the lack of accuracy in the coefficient matrix and right-hand-side vectors did not result in large changes in the solution, indicating that the steps taken to address potential ill-conditioning of the matrix were effective.

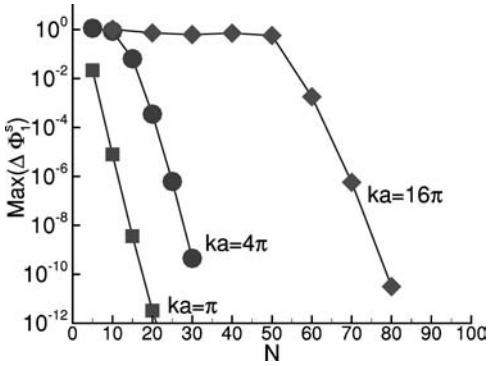


FIG. 2. Numerical convergence behavior for two-cylinder configuration with various values of ka. $\Delta\Phi_1^s$ represents the maximum change for the surface scattered field between solutions generated with N and $N+\Delta N$ terms.

Next, the issue of number of terms needed to obtain a converged solution is examined. For plane-wave scattering from symmetric, two-cylinder configurations, Young and Bertrand[7] and Decanini *et al.*[23] give empirical expressions for the value of N as

$$N=2ka+Q, \quad Q=\begin{cases}5-10, & ka\leqslant 1\\ 0, & ka>1,\end{cases} \tag{29}$$

and

$$N=\max\{8,[ka+4(ka)^{1/3}+1]\}, \tag{30}$$

respectively. Here, the number of terms required for a given level of accuracy is evaluated for two particular configurations. The first consists of three identical cylinders located at the vertices of an equilateral triangle with sides of length 3 and a line source located at the centroid of the configuration. The second consists of two identical cylinders separated by a distance of 8, with a spatially Gaussian source given by Eq. (27), where $d=25\ln 2$ located midway between them. To check convergence, the scattered field at the surface of one of the cylinders was tracked as the value of N was increased by steps of ΔN. The maximum difference between the scattered field obtained with N coefficients and the scattered field obtained with $N+\Delta N$ is plotted in Fig. 2 for the two-cylinder configuration and Fig. 3 for the three-cylinder configuration for various values of ka. Here, $\Delta N=5$ for all cases except the $ka=16\pi$ case, in which $\Delta N=10$. Examination of these figures shows that the formula given by Eq. (30) matches very well with the number of terms required to obtain a converged solution, here defined as truncation errors of less than 10^{-6}. However, the simpler formula given by Eq. (29) underpredicts the number of coefficients required at the lowest value of $ka=\pi$ and overpredicts N at the highest value of $ka=16\pi$. Thus, Eq. (30) was used to determine an initial value of N to use for subsequent calculations, although slightly higher values of N were also used to verify convergence.

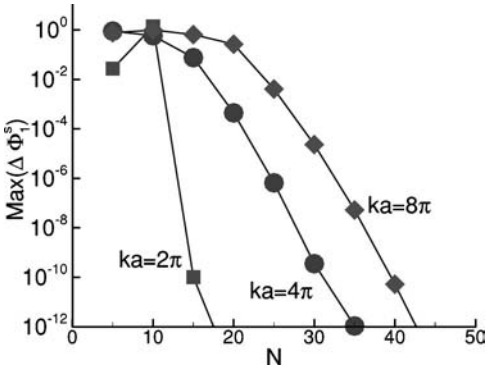


FIG. 3. Numerical convergence behavior for three-cylinder configuration with various values of ka. $\Delta\Phi_1^s$ represents the maximum change for the surface scattered field between solutions generated with N and $N+\Delta N$ terms.

Various validation checks were performed on the multi-scattering formulation presented here in order to verify its consistency. For the single-cylinder scattering case, the scattering coefficients from this formulation were found to degenerate to those presented by Morris[5] in the rigid-body limit. It was also verified that the formulation yielded the same results for configurations that were different numerically but nearly identical acoustically, i.e., configurations that have a different number of cylinders but, because of the size and/or locations of one or more of the cylinders, appear from an acoustic standpoint to be the same. Results from two such cases are shown in Fig. 4. In the first case, computed with $k=8\pi$, an initially symmetric two-cylinder configuration is degenerated into a one-cylinder configuration by fixing the location and size of one of the cylinders and either incrementally increasing the distance between the source and the other cylinder towards infinity or incrementally decreasing its radius towards zero. The second case, with $k=2\pi$, involved the degeneration of an initially symmetric four-cylinder configuration into a symmetric two-cylinder configuration by fixing the two cylinders on opposite sides of the source and either increasing the distances between the source and the other two cylinders in opposite directions, or decreasing the radii of these two cylinders. Figure 4 plots the scattered field in the backscatter direction (along the line connecting the source and cylinder centers) on the surface of a fixed cylinder versus the parameter n, which controls either the distance between the cylinders being acoustically removed and the source via the formula $L=2^n$, or radii of the removed cylinders via $a=(\frac{1}{2})^n$. For all cases, the results obtained by acoustically removing one or more cylinders from consideration match those obtained if the removed cylinders were never considered. Additional validation was also provided by comparison with numerical results obtained for various two- and three-cylinder configurations via a finite-difference time-domain computational aeroacoustics code.[17,18]

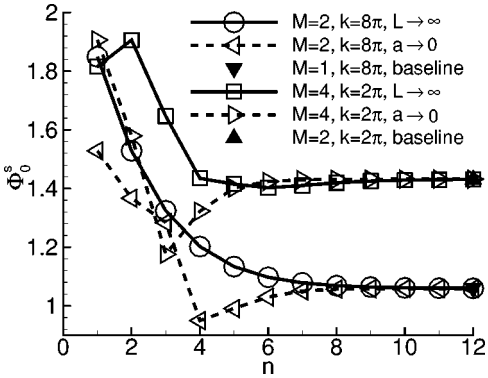


FIG. 4. Scattered field behavior in the backscatter direction for two- and four-cylinder configurations as cylinder(s) are acoustically removed from problem either by increasing the distance between the cylinder(s) and the source ($L\to\infty$) or decreasing their radii ($a\to 0$).

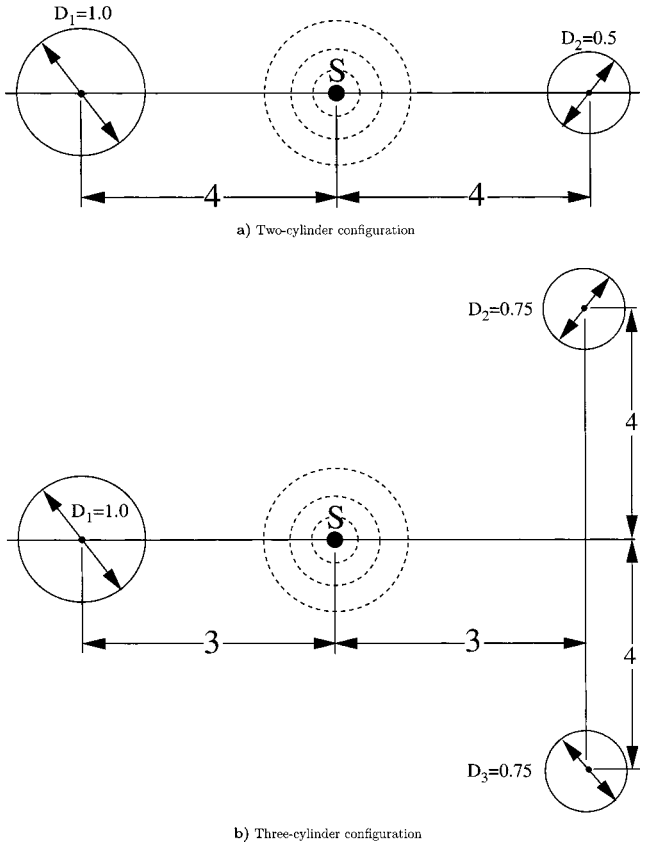


FIG. 5. Benchmark acoustic scattering geometries. (a) Two-cylinder configuration. (b) Three-cylinder configuration.

B. Two- and three-cylinder benchmark configurations

Two specific cases are considered here to provide benchmark results for use in the validation of computational aeroacoustic methods developed for complex geometries.[19] These consist of a two-cylinder configuration and a three-cylinder configuration as shown in Fig. 5. Both cases are symmetric about the centerline axis (the axis connecting the centers of the left-most cylinder and the source), but are asymmetric about the perpendicular axis. Each case was solved with $k=8\pi$ using both line and spatially distributed Gaussian sources, where $d=25\ln 2$ for the latter. Results presented here include the scattered and total root-mean-squared (rms) fields on the surface of the cylinders and the total rms fields along the centerline. These are shown in Figs. 6 through 9 for the two-cylinder configuration, and in Figs. 10 through 13 for the three-cylinder configuration. It is observed that the

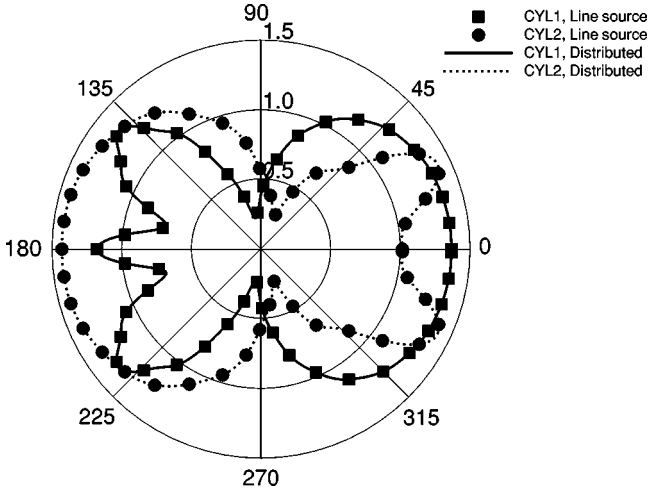


FIG. 6. Scattered field on surfaces of the cylinders for the two-cylinder configuration shown in Fig. 5(a).

line source and spatially distributed Gaussian source produce nearly identical results everywhere except in the immediate vicinity of the source. In this region, the total fields generated by the spatially distributed source begin increasing further away from the source center compared to the field generated by the point source. However, the total field associated with the distributed source remains finite everywhere, while the total field for the line source becomes infinite at the source center.

For the two-cylinder configuration, the scattered field and total rms field on the surface of each cylinder are shown in Figs. 6 and 7, respectively. They are characterized by a smooth variation on the insonified faces of the cylinders. The fields on the quiet sides of the cylinders exhibit more variation, although the magnitude of the total field in Fig. 7 is considerably less on the quiet faces compared to the insonified faces due to the higher incident field magnitudes there. Due to its larger radius, the cylinder to the left of the source exhibits more variation, both in terms of the number of peaks and nulls in the field strength as well as the range between them, over its quiet face compared to the smaller, right-hand cylinder. However, the peak total field on the surface at the point directly facing the source is slightly higher for the smaller cylinder due to the greater scattered field it is exposed to from the larger cylinder. The total rms field calculated along the centerline is shown in Figs. 8 and 9. In both figures, the field along the centerline between the source and the cylinders is highly oscillatory, with a smooth variation in peak and null magnitudes. The behavior of the total field in the shadowed regions is characterized by a slight rise in magnitude near the cylinders followed by a gradual decline as the distance from the cylinder is increased.

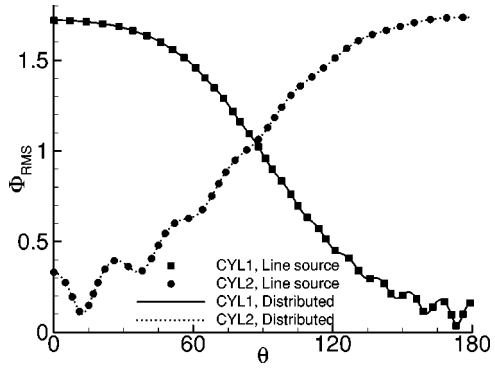


FIG. 7. Total rms field on surfaces of the cylinders for the two-cylinder configuration shown in Fig. 5(a).

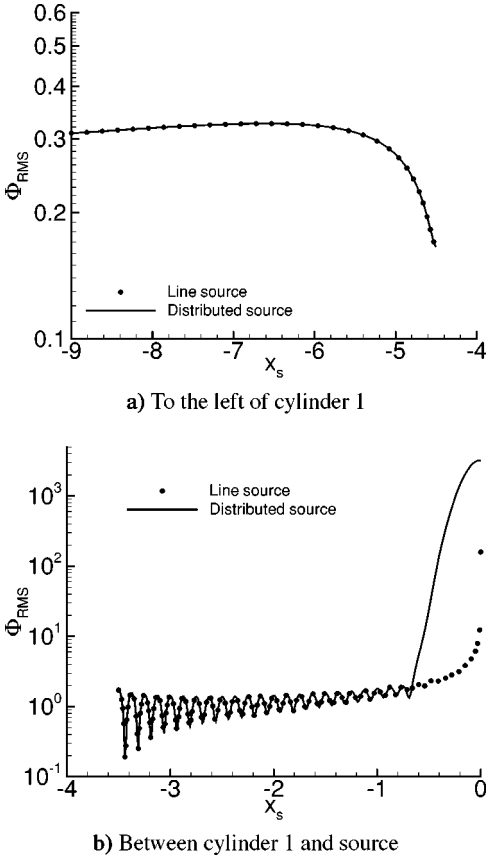
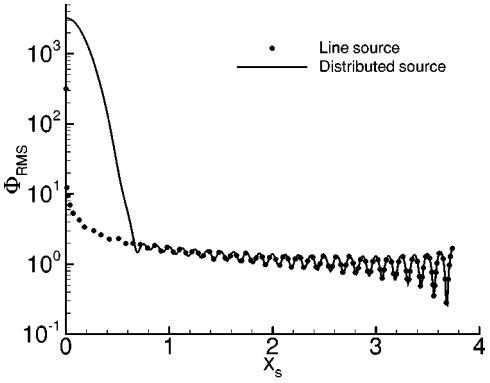
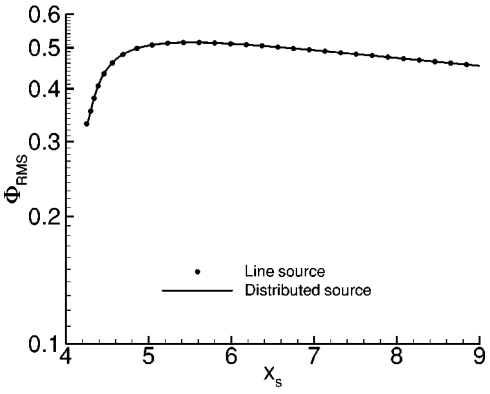


FIG. 8. Total rms field on the centerline to the left of the source for the two-cylinder configuration shown in Fig. 5(a). (a) To the left of cylinder 1. (b) Between cylinder 1 and source.

The scattered and total rms fields on the surfaces of the left- and top-right cylinders in the three-cylinder configuration are shown in Figs. 10 and 11. The surface fields for the three-cylinder case show considerably more variation with angular position than was seen in the two-cylinder case, both on the insonified faces (between approximately 143 and 323 degrees for the top-right cylinder) and on the quiet faces. Because the cylinders to the right of the source experience the incoming incident and scattered fields from different directions, no angular symmetry exists for these two cylinders. The centerline mean-squared total field is shown in Figs. 12 and 13 and also exhibits more complexity than was seen in the two-cylinder case, as the magnitudes of the peaks and nulls themselves vary in an oscillatory manner between the left cylinder and the source. Also, the field in the shadow region, while still diminishing at larger distances from the cylinder, shows additional variation due to the more complex scattering geometry that was not seen in the two-cylinder case.



a) Between source and cylinder 2



b) To the right of cylinder 2

FIG. 9. Total rms field on the centerline to the right of the source for the two-cylinder configuration shown in Fig. 5(a). (a) Between source and cylinder 2. (b) To the right of cylinder 2.

IV. CONCLUSIONS

An analytic method for evaluating the scattered fields created by a collection of rigid, two-dimensional, circular cylinders when exposed to an incident field generated by two types of axisymmetric sources has been developed. The source types considered here include a cylindrical line source and a source with a Gaussian spatial distribution. The approach uses a Hankel transform method to determine the incident field, and separation of variables to find the scattered fields from each cylinder in the collection. The unknown scattering coefficients are then determined through the solution of a linear system of equations obtained from the simultaneous application of boundary conditions on the surfaces of all cylinders. For a line source, analytic expressions for all terms in this system of equations may be obtained, while numerical evaluation of some integrals is required for the spatially distributed source. Various two-, three-, and four-cylinder symmetric configurations were examined for validation purposes and to investigate the number of coefficients required to obtain a converged solution. Also, results are presented for asymmetric two- and three-cylinder con-

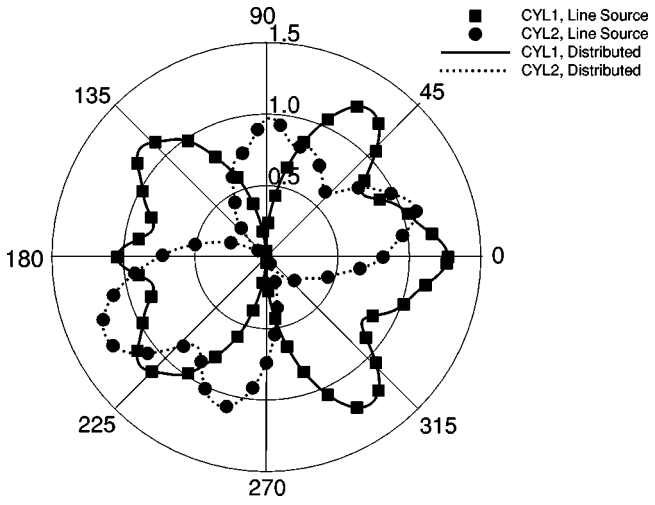


FIG. 10. Scattered field on surfaces of the left and top-right cylinders for the three-cylinder configuration shown in Fig. 5(b).

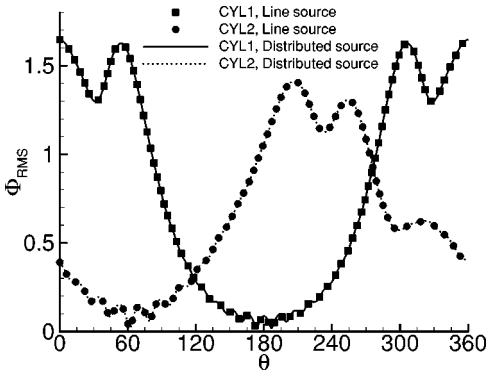
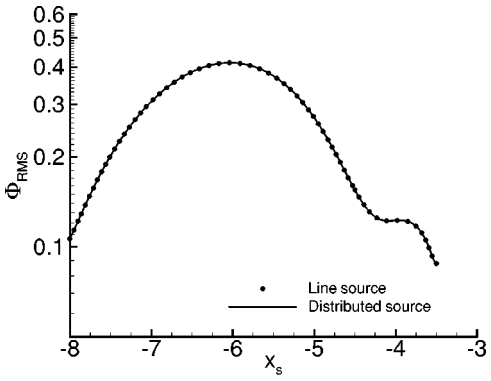
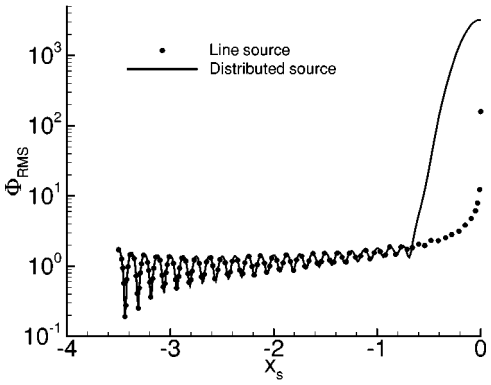


FIG. 11. Total rms field on surfaces of the left- and top-right cylinders for the two-cylinder configuration shown in Fig. 5(b).



a) To the left of cylinder 1



b) Between cylinder 1 and the source

FIG. 12. Total rms field on the centerline to the left of the source for the three-cylinder configuration shown in Fig. 5(b). (a) To the left of cylinder 1. (b) Between cylinder 1 and the source.

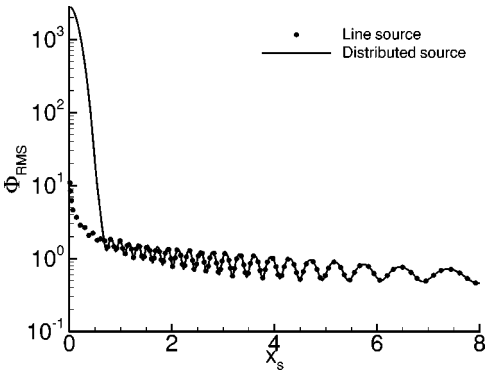


FIG. 13. Total rms field on the centerline to the right of the source for the three-cylinder configuration shown in Fig. 5(b).

figurations that have proved useful for the validation of computational aeroacoustic codes developed for complex geometries.

ACKNOWLEDGMENTS

The author wishes to thank Dr. Phillip Morris for providing FORTRAN versions of his numerical integration routine for examination. The author would also like to thank Dr. Donald Rizzetta and Dr. Miguel Visbal as well as the referees of this work for their reviews of the manuscript and helpful comments provided.

[1] P. Morse and K. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
[2] C. A. Balanis, *Advanced Engineering Electromagnetics*, 1st ed. (Wiley, New York, 1989).
[3] E. Shenderov, ''Diffraction of a cylindrical sound wave by a cylinder,'' Sov. Phys. Acoust. **7**, 293–296 (1962).
[4] K. A. Kurbatskii, ''Analytical Solutions of the Category 1, Benchmark Problems 1 and 2,'' Second Computational Aeroacoustic (CAA) Workshop on Benchmark Problems, NASA CP-3352, 1997, pp. 9–14.
[5] P. J. Morris, ''The scattering of sound from a spatially distributed symmetric cylindrical source by a circular cylinder,'' J. Acoust. Soc. Am. **97**, 2651–2656 (1995).
[6] V. Twersky, ''Multiple scattering of radiation by an arbitrary configuration of parallel cylinders,'' J. Acoust. Soc. Am. **24**, 42–46 (1952).
[7] J. Young and J. Bertrand, ''Multiple scattering by two cylinders,'' J. Acoust. Soc. Am. **58**, 1190–1195 (1975).
[8] B. Peterson and S. Ström, ''Matrix formulation of acoustic scattering from an arbitrary number of scatterers,'' J. Acoust. Soc. Am. **56**, 771–780 (1974).
[9] P. Waterman, ''New formulation of acoustic scattering,'' J. Acoust. Soc. Am. **45**, 1417–1429 (1969).
[10] A. Elsherbeni, ''A comparative study of two-dimensional multiple scattering techniques,'' Radio Sci. **29**, 1023–1033 (1994).
[11] J. Vasseur, P. Deymier, B. Chenni, B. Djafari-Rouhani, L. Dobrzynski, and D. Prevost, ''Experimental and theoretical evidence for the existence of absolute acoustic band gaps in two-dimensional solid phononic crystals,'' Phys. Rev. Lett. **86**, 3012–3015 (2001).
[12] M. Kafesaki and E. Economou, ''Multiple-scattering theory for three-dimensional periodic acoustic composites,'' Phys. Rev. B **60**, 11993–12001 (1999).
[13] M. Kafesaki, R. Penciu, and E. Economou, ''Air bubbles in water: A strongly multiple scattering medium for acoustic waves,'' Phys. Rev. Lett. **84**, 6050–6053 (2000).
[14] I. Psarobas, N. Stefanou, and A. Modinos, ''Scattering of elastic waves by periodic arrays of spherical bodies,'' Phys. Rev. B **62**, 278–291 (2000).
[15] S. Yang, J. Page, Z. Liu, M. Cowan, C. Chan, and P. Sheng, ''Ultrasound tunneling through 3D phononic crystals,'' Phys. Rev. Lett. **88**, 104301 (2002).
[16] B. Gupta and Z. Ye, ''Localization of classical waves in two-dimensional random media: A comparison between the analytic theory and exact numerical simulation,'' Phys. Rev. E **67**, 036606 (2003).
[17] S. Sherer and M. Visbal, ''Computational study of acoustic scattering from multiple bodies using a high-order overset grid approach,'' AIAA Paper 2003-3203, 9th AIAA/CEAS Aeroacoustics Conference, Hilton Head, SC, June 2003.
[18] S. E. Sherer, ''Further analysis of high-order overset grid method with applications,'' AIAA Paper 2003-3839, AIAA 16th Computational Fluid Dynamics Conference, Orlando, FL, June 2003.
[19] Fourth Computational Aeroacoustic (CAA) Workshop on Benchmark Problems, NASA Conference Proceeding, 2004.
[20] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, 5th ed. (Dover, New York, 1968).
[21] S. Wolfram, *The Mathematica Book* (Cambridge University Press, New York, 1999).
[22] G. Strang, *Linear Algebra and Its Applications*, 3rd ed. (Harcourt Brace, New York, 1988).
[23] Y. Decanini, A. Folacci, P. Gabrielli, and J.-L. Rossi, ''Algebraic aspects of multiple scattering by two parallel cylinders: Classification and physical interpretation of scattering resonances,'' J. Sound Vib. **221**, 785–804 (1999).

Coherence function and mean field of plane and spherical sound waves propagating through inhomogeneous anisotropic turbulence

Vladimir E. Ostashev
NOAA/Environmental Technology Laboratory, 325 Broadway, Boulder, Colorado 80305 and Physics Department, New Mexico State University, Las Cruces, New Mexico

D. Keith Wilson
U.S. Army Cold Regions Research and Engineering Laboratory, 72 Lyme Road, Hanover, New Hampshire 03755

(Received 1 July 2003; accepted for publication 17 November 2003)

Inhomogeneity and anisotropy are intrinsic characteristics of daytime and nighttime atmospheric turbulence. For example, turbulent eddies are often stretched in the direction of the mean wind, and the turbulence statistics depends on the height above the ground. Recent studies have shown that the log-amplitude and phase fluctuations of plane and spherical sound waves are significantly affected by turbulence inhomogeneity and anisotropy. The present paper is devoted to studies of the mean sound field and the coherence functions of plane and spherical sound waves propagating through inhomogeneous anisotropic turbulence with temperature and velocity fluctuations. These statistical moments of a sound field are important in many practical applications, e.g., for source detection, ranging, and recognition. Formulas are derived for the mean sound field and coherence function of initially arbitrary waveform. Using the latter formula, we also obtained formulas for the coherence functions of plane and spherical sound waves. All these formulas coincide with those known in the literature for two limiting cases: homogeneous isotropic turbulence with temperature and wind velocity fluctuations, and inhomogeneous anisotropic turbulence with temperature fluctuations only. Using the formulas obtained, we have numerically shown that turbulence inhomogeneity significantly affects the coherence functions of plane and spherical sound waves. [DOI: 10.1121/1.1639339]

PACS numbers: 43.20.Fn, 43.20.Bi, 43.28.Gq [MO] Pages: 497–506

I. INTRODUCTION

The coherence function Γ and the mean field $\langle p \rangle$ are important statistical characteristics of a sound wave propagating in a turbulent atmosphere.[1,2] The values of Γ and $\langle p \rangle$ are needed for many practical applications in modern atmospheric acoustics. First, analytical formulas for Γ and $\langle p \rangle$ are used in assessing performance of acoustic arrays for source detection, ranging, and recognition.[3,4] Second, a comparison between an analytical expression for Γ and measurements of the coherence function can be used for acoustic remote sensing of statistical characteristics of atmospheric turbulence.[5] Third, theoretical estimates of Γ are important for designing infrasound arrays for the Comprehensive Test Ban Treaty: The distance between microphones in these arrays should be less than the coherence radius of an infrasound wave. These and other applications in atmospheric acoustics make the coherence function and the mean field one of the most important statistical moments of a sound field. Also note that calculations of Γ and $\langle p \rangle$ are of interest for a relatively new field in physics: waves in random media,[6–13] which has been rapidly developing since the middle of the last century.

In this paper, we will consider line-of-sight wave propagation in a turbulent medium. For this case, formulas for the coherence function Γ and the mean field $\langle p \rangle$ were first derived for an electromagnetic wave propagating in a turbulent atmosphere.[8,10] These formulas can also be used to describe Γ and $\langle p \rangle$ for the case of a sound wave propagation in an atmosphere with temperature fluctuations.

However, wind velocity fluctuations are always present in the atmosphere and affect significantly Γ and $\langle p \rangle$. Formulas for Γ and $\langle p \rangle$ for plane and spherical sound waves propagating through homogeneous isotropic turbulence with temperature and wind velocity fluctuations were summarized in Ref. 13. Note that temperature fluctuations constitute a scalar random field, while velocity fluctuaitions constitute a vector random field; the spectral tensor of velocity fluctuations is anisotropic even for homogeneous isotropic turbulence. As a result, the effects of velocity fluctuations on Γ and $\langle p \rangle$ are qualitatively different from those of temperature fluctuations.

The main goal of the present paper is to generalize the formulas for Γ and $\langle p \rangle$ obtained for the case of sound propagation through homogeneous isotropic turbulence with temperature and velocity fluctuations to the case of inhomogeneous anisotropic turbulence.

Daytime and nighttime atmospheric turbulence is intrinsically inhomogeneous and anisotropic. Indeed, statistical characteristics of atmospheric turbulence (e.g., the variances of wind velocity fluctuations along different directions) significantly depend on the height above the ground. This manifests statistical inhomogeneity of atmospheric turbulence. Furthermore, these statistical characteristics are different in different directions (e.g., the variances of velocity fluctua-

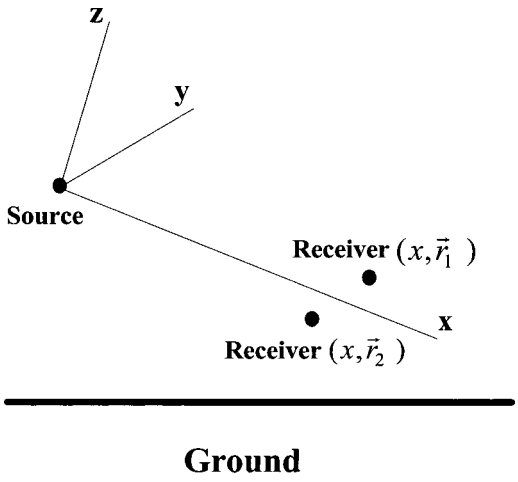


FIG. 1. Geometry of line-of-sight sound propagation.

tions are different along wind, cross-wind, and in the vertical direction), manifesting statistical anisotropy of turbulence. It was shown in Ref. 14 that anisotropy of turbulence affects significantly sonic boom propagation. Furthermore, recent studies have shown[15,16] that turbulence inhomogeneity and anisotropy affect significantly the variances and correlation functions of log-amplitude and phase fluctuations. Therefore, the derivation and analysis of formulas for Γ and $\langle p\rangle$ for plane and spherical sound waves for the case of inhomogeneous anisotropic turbulence, which are presented in this paper, are important steps in developing theories of sound propagation through a turbulent atmosphere. Note that the effects of turbulence anisotropy on Γ and $\langle p\rangle$ for the case of a plane wave propagation were considered elsewhere.[17] Some results obtained in the present paper were reported at meetings.[18,19]

The paper is organized as follows. Section II deals with starting equations of our analysis. In Sec. III, formulas for the coherence functions of plane and spherical sound waves are obtained. The mean sound field is considered in Sec. IV. The effects of turbulence inhomogeneity on the coherence functions of plane and spherical sound waves are studied numerically in Sec. V. The results obtained in the paper are summarized in Sec. VI.

II. GEOMETRY OF THE PROBLEM AND STARTING EQUATIONS

A. Geometry

The geometry of the problem is shown in Fig. 1. A point source is located at the origin of the coordinate system x, y, z. The x axis is in the direction close to the geometrical center of two receivers, which are used to measure the coherence function Γ. The receivers are located in the plane $x=\text{const}$ at the points $(x,\mathbf{r}_1)$ and $(x,\mathbf{r}_2)$. Here, $\mathbf{r}=(y,z)$ are the transverse coordinates. Temperature $\tilde{T}(x,\mathbf{r})$ and velocity $\mathbf{v}(x,\mathbf{r})$ fluctuations fill the half-space $x>0$.

We will also consider the case when a plane sound wave is incident on this half-space from the left in the direction of the x axis. For both plane wave and point source, the initial sound field $p_0(\mathbf{r})$, which is not disturbed by random inhomogeneities in a medium, is given in the plane $x=0$.

Note that the x axis might or might not be parallel to the ground. Furthermore, the receivers might be located on or above the ground. In either case, a sound wave reflected from the ground (if there is any) is neglected.

B. Starting equations

The transverse coherence function of a sound field p is defined as

$$\Gamma(x;\mathbf{r}_1,\mathbf{r}_2)=\langle p(x,\mathbf{r}_1)p^*(x,\mathbf{r}_2)\rangle. \tag{1}$$

Here, the brackets $\langle\ \rangle$ denote ensemble average.

A closed equation for Γ is derived in Ref. 20 for the case of sound propagation in a turbulent, refractive atmosphere near an impedance boundary. This equation is derived starting from the parabolic equation for sound waves and using the Markov approximation which is widely used in waves in random media.[10] The equation is valid for sound propagation through inhomogeneous anisotropic turbulence with temperature and velocity fluctuations. For the case of line-of-sight sound propagation, considered in the present paper, the closed equation for Γ simplifies and takes the form

$$\frac{\partial\Gamma(x;\mathbf{r}_1,\mathbf{r}_2)}{\partial x}-\frac{i}{2k}\left(\frac{\partial^2}{\partial\mathbf{r}_1^2}-\frac{\partial^2}{\partial\mathbf{r}_2^2}\right)\Gamma+\frac{k^2}{8}M(x;\mathbf{r}_1,\mathbf{r}_2)\Gamma=0. \tag{2}$$

Here $x>0$, k is a reference sound wave number, and the function M is given by

$$M(x;\mathbf{r}_1,\mathbf{r}_2)=b_{\text{eff}}(x;\mathbf{r}_1,\mathbf{r}_1)+b_{\text{eff}}(x;\mathbf{r}_2,\mathbf{r}_2)-2b_{\text{eff}}(x;\mathbf{r}_1,\mathbf{r}_2), \tag{3}$$

where b_{eff} is the transverse effective correlation function, defined by the following formula:

$$\left\langle\left(\frac{\tilde{T}(x_1,\mathbf{r}_1)}{T_0}+\frac{2v_x(x_1,\mathbf{r}_1)}{c_0}\right)\left(\frac{\tilde{T}(x_2,\mathbf{r}_2)}{T_0}+\frac{2v_x(x_2,\mathbf{r}_2)}{c_0}\right)\right\rangle=\delta(x_1-x_2)b_{\text{eff}}(x;\mathbf{r}_1,\mathbf{r}_2). \tag{4}$$

In Eq. (4), T_0 and c_0 are the mean values of temperature and adiabatic sound speed, v_x is the component of wind velocity in the direction of the x axis, and δ is the delta-function.

The initial condition to Eq. (2) is formulated in the plane $x=0$:

$$\Gamma_0(\mathbf{r}_1,\mathbf{r}_2)=\Gamma(x=0;\mathbf{r}_1,\mathbf{r}_2)=p_0(\mathbf{r}_1)p_0^*(\mathbf{r}_2). \tag{5}$$

Equation (2) for the coherence function Γ with initial condition (5) cannot be solved analytically for the case of inhomogeneous turbulence. In the next section, we present an approximate solution of this equation.

III. COHERENCE FUNCTION

A. Approximate formula for the coherence function

In Eqs. (1)–(5), it is convenient to introduce new transverse coordinates:

$$\mathbf{r}_-=\mathbf{r}_1-\mathbf{r}_2,\quad \mathbf{r}_+=(\mathbf{r}_1+\mathbf{r}_2)/2. \tag{6}$$

Then, $\Gamma(x;\mathbf{r}_1,\mathbf{r}_2)$, $\Gamma_0(\mathbf{r}_1,\mathbf{r}_2)$, $M(x;\mathbf{r}_1,\mathbf{r}_2)$, and $b_{\text{eff}}(x;\mathbf{r}_1,\mathbf{r}_2)$ are written as $\Gamma(x;\mathbf{r}_+ +\mathbf{r}_-/2,\mathbf{r}_+-\mathbf{r}_-/2)$, $\Gamma_0(\mathbf{r}_+ +\mathbf{r}_-/2,\mathbf{r}_+$

$-\mathbf{r}_-/2)$, $M(x;\mathbf{r}_+ + \mathbf{r}_-/2, \mathbf{r}_+ - \mathbf{r}_-/2)$, and $b_{\text{eff}}(x;\mathbf{r}_+ + \mathbf{r}_-/2, \mathbf{r}_+ - \mathbf{r}_-/2)$. In what follows (except Appendix B), the latter four functions will be denoted as $\Gamma(x;\mathbf{r}_+,\mathbf{r}_-)$, $\Gamma_0(\mathbf{r}_+,\mathbf{r}_-)$, $M(x;\mathbf{r}_+,\mathbf{r}_-)$, and $b_{\text{eff}}(x;\mathbf{r}_+,\mathbf{r}_-)$.

With these notations, Eq. (2) takes the form

$$\frac{\partial \Gamma(x;\mathbf{r}_+,\mathbf{r}_-)}{\partial x} - \frac{i}{k}\frac{\partial}{\partial \mathbf{r}_-}\cdot\frac{\partial}{\partial \mathbf{r}_+}\Gamma + \frac{k^2}{8}M(x;\mathbf{r}_+,\mathbf{r}_-)\Gamma = 0, \tag{7}$$

where

$$M(x;\mathbf{r}_+,\mathbf{r}_-) = b_{\text{eff}}(x;\mathbf{r}_+ + \mathbf{r}_-/2,0) + b_{\text{eff}}(x;\mathbf{r}_+ - \mathbf{r}_-/2,0) - 2b_{\text{eff}}(x;\mathbf{r}_+,\mathbf{r}_-). \tag{8}$$

In Ref. 15, the function $b_{\text{eff}}(x;\mathbf{r}_+,\mathbf{r}_-)$ is expressed in terms of the effective three-dimensional spectral density Φ_{eff} of medium inhomogeneities:

$$b_{\text{eff}}(x;\mathbf{r}_+,\mathbf{r}_-) = 2\pi\int d\mathbf{K}_\perp\, e^{i\mathbf{K}_\perp\cdot\mathbf{r}_-}\Phi_{\text{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp), \tag{9}$$

where Φ_{eff} is given by

$$\Phi_{\text{eff}}(x,\mathbf{r}_+;K_x,\mathbf{K}_\perp) = \frac{\Phi_T(x,\mathbf{r}_+;K_x,\mathbf{K}_\perp)}{T_0^2} + \frac{4\Phi_{xx}(x,\mathbf{r}_+;K_x,\mathbf{K}_\perp)}{c_0^2} + \frac{2}{T_0 c_0}\times[\Phi_{Tx}(x,\mathbf{r}_+;K_x,\mathbf{K}_\perp) + \Phi_{Tx}(x,\mathbf{r}_+;-K_x,-\mathbf{K}_\perp)]. \tag{10}$$

Here, Φ_T, Φ_{xx}, and Φ_{Tx} are the three-dimensional spectral densities of fluctuations in temperature, velocity, and cross-correlation between temperature and velocity. Furthermore, in Eqs. (9) and (10), K_x and $\mathbf{K}_\perp$ are the components of the turbulence wave vector $\mathbf{K}=(K_x,\mathbf{K}_\perp)$. The dependence of the spectral densities Φ_{eff}, Φ_T, Φ_{xx}, and Φ_{Tx} on x and $\mathbf{r}_+$ is due to inhomogeneity of turbulence. For homogeneous turbulence, these spectral densities do not depend on x and $\mathbf{r}_+$, and Eq. (10) (with $\Phi_{Tx}=0$) coincides with Eq. (7.43) from Ref. 13.

Substituting the value of $b_{\text{eff}}(x;\mathbf{r}_+,\mathbf{r}_-)$ given by Eq. (9) into Eq. (8), we obtain

$$M(x;\mathbf{r}_+,\mathbf{r}_-) = 2\pi\int d\mathbf{K}_\perp[\Phi_{\text{eff}}(x,\mathbf{r}_+ + \mathbf{r}_-/2;0,\mathbf{K}_\perp) + \Phi_{\text{eff}}(x,\mathbf{r}_+ - \mathbf{r}_-/2;0,\mathbf{K}_\perp) - 2e^{i\mathbf{K}_\perp\cdot\mathbf{r}_-}\Phi_{\text{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp)]. \tag{11}$$

Equation (7) is equivalent to Eq. (2) and cannot be solved analytically in a general case. To approximately solve Eq. (7), let us consider the scale a of variations in the effective spectral density $\Phi_{\text{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp)$ with respect to the coordinate $\mathbf{r}_+$. For the case of homogeneous turbulence, Φ_{eff} does not depend on $\mathbf{r}_+$ (and x) so that $a\to\infty$. For another particular case, when the outer length scale of turbulence and the variances of temperature and velocity fluctuations depend only on the height above the ground, the value of a depends on the direction of the sound propagation path. If this path is in the vertical direction, $a\to\infty$. On the other hand, if the path is close to horizontal, a is of order of a vertical scale of variations in the outer length scale and variances with height.

In what follows, we assume that the scale a is greater than the size of the first Fresnel zone $\sqrt{x/k}$. Using this assumption, Eq. (7) is solved in Appendix A. As a result, we obtain a desired formula for the coherence function of a sound wave propagating through inhomogeneous anisotropic turbulence with temperature and velocity fluctuations:

$$\Gamma(x;\mathbf{r}_+,\mathbf{r}_-) = \frac{k^2}{4\pi^2x^2}\int d\mathbf{r}'_-\int d\mathbf{r}'_+\,\Gamma_0(\mathbf{r}'_+,\mathbf{r}'_-)\times\exp\left[i(k/x)(\mathbf{r}_- - \mathbf{r}'_-)\cdot(\mathbf{r}_+ - \mathbf{r}'_+) - \frac{k^2}{8}\times\int_0^x dx'\,\tilde{M}(x';0,\mathbf{r}_- x'/x + \mathbf{r}'_-(1-x'/x))\right]. \tag{12}$$

Here, the function $\tilde{M}$ is given by Eq. (A12). Equation (12) is valid for an arbitrary sound field $p_0(\mathbf{r})$ in the plane $x=0$. Formulas for the coherence functions of plane and spherical sound waves are presented in the next two subsections.

B. Plane wave

For the case of plane wave propagation, $p_0(\mathbf{r})=A_0$, where A_0 is the amplitude of a sound field in the plane $x=0$. Therefore, $\Gamma_0(\mathbf{r}'_+,\mathbf{r}'_-)=|A_0|^2$.

We substitute the value of Γ_0 into Eq. (12). Then, the integral over $\mathbf{r}'_+$ is equal to $4\pi^2\delta((k/x)(\mathbf{r}_- - \mathbf{r}'_-))$. Therefore, the remaining integral over $\mathbf{r}'_-$ can be easily evaluated. As a result, we have

$$\Gamma^p(x;\mathbf{r}_+,\mathbf{r}_-) = |A_0|^2\exp\left[-\frac{k^2x}{8}\int_0^1 d\eta\,\tilde{M}(\eta x;0,\mathbf{r}_-)\right]. \tag{13}$$

Substituting the value of $\tilde{M}$ given by Eq. (A12) into Eq. (13), we get a final formula for the coherence function of a plane sound wave:

$$\Gamma^p(x;\mathbf{r}_+,\mathbf{r}_-) = |A_0|^2\exp\left[-\frac{\pi k^2x}{2}\int_0^1 d\eta\int d\mathbf{K}_\perp\times(1-e^{i\mathbf{K}_\perp\cdot\mathbf{r}_-})\Phi_{\text{eff}}(\eta x,0;0,\mathbf{K}_\perp)\right]. \tag{14}$$

This equation is valid for sound propagation through inhomogeneous anisotropic turbulence with temperature and velocity fluctuations. It generalizes equations known in the literature. First, for homogeneous isotropic turbulence, Eq. (14) coincides with Eq. (7.63) from Ref. 13. Second, for homogeneous anisotropic turbulence, it coincides with Eq. (15) from Ref. 17. [Note that Eq. (15) from Ref. 17 was adopted from the literature rather than derived as in Appendix A.] Third, for inhomogeneous anisotropic turbulence with temperature fluctuations only, Eq. (14) is essentially the same as Eq. (3.21) from Ref. 21.

C. Spherical wave

In the approximation of a parabolic equation, a sound field in the plane $x=0$ due to a point source located at the

origin of the coordinate system is given by $p_0(\mathbf{r}) = (i/2k)\delta(\mathbf{r})$. Therefore, in this case, the function $\Gamma_0(\mathbf{r}'_+, \mathbf{r}'_-) = (1/4k^2)\delta(\mathbf{r}'_+ + \mathbf{r}'_-/2)\delta(\mathbf{r}'_+ - \mathbf{r}'_-/2)$. Substituting this formula into Eq. (12) and calculating the integrals with respect to $\mathbf{r}'_+$ and $\mathbf{r}'_-$, we have

$$\Gamma^s(x;\mathbf{r}_+,\mathbf{r}_-) = \frac{1}{(4\pi x)^2}\exp\left[\frac{ik\mathbf{r}_+\cdot\mathbf{r}_-}{x} - \frac{k^2 x}{8}\int_0^1 d\eta\, \tilde{M}(\eta x;0,\eta\mathbf{r}_-)\right]. \tag{15}$$

Replacing the function $\tilde{M}$ by its value given by Eq. (A12), we obtain a final formula for the coherence function of a spherical sound wave propagating through inhomogeneous anisotropic turbulence with temperature and velocity fluctuations:

$$\Gamma^s(x;\mathbf{r}_+,\mathbf{r}_-) = \frac{1}{(4\pi x)^2}\exp\left[\frac{ik\mathbf{r}_+\cdot\mathbf{r}_-}{x} - \frac{\pi k^2 x}{2}\int_0^1 d\eta \int d\mathbf{K}_\perp \times (1 - e^{i\eta\mathbf{K}_\perp\cdot\mathbf{r}_-})\Phi_{\text{eff}}(\eta x,0;0,\mathbf{K}_\perp)\right]. \tag{16}$$

This formula for Γ^s also generalizes those known in the literature. For the case of homogeneous isotropic turbulence, Eq. (16) is the same as Eq. (7.71) from Ref. 13. For inhomogeneous anisotropic turbulence with temperature fluctuations only, Eq. (16) can be obtained from Eq. (3.28) in Ref. 21.

IV. MEAN SOUND FIELD

Probably the simplest way to derive a formula for the mean sound field is to set $\mathbf{r}_- \to \infty$ in Eq. (14). Then, the left-hand side of the resulting equation can be written as a product of the mean sound field and its complex conjugate: $\langle p\rangle\langle p^*\rangle$. On the right-hand side, the term $\exp(i\mathbf{K}_\perp\cdot\mathbf{r}_-)$ can be omitted since it oscillates rapidly and does not contribute to the integral over $\mathbf{K}_\perp$. As a result, we obtain the following formula for the mean sound field propagating through inhomogeneous anisotropic turbulence with temperature and velocity fluctuations:

$$\langle p(x,\mathbf{r})\rangle = p_0(x,\mathbf{r}) \times \exp\left[-\frac{\pi k^2 x}{4}\int_0^1 d\eta \int d\mathbf{K}_\perp\, \Phi_{\text{eff}}(\eta x,0;0,\mathbf{K}_\perp)\right]. \tag{17}$$

Here, $p_0(x,\mathbf{r})$ is the sound field in a medium without turbulence.

Note that Eq. (17) can also be obtained starting from Eq. (16) for Γ^s and using the approach above. A more rigorous derivation of Eq. (17) is given in Appendix B.

Particular cases of Eq. (17) can be found in the literature. For the case of homogeneous anisotropic turbulence, Eq. (17) coincides with Eq. (14) in Ref. 17 and for homogeneous isotropic turbulence with Eq. (7.60) from Ref. 13.

V. CALCULATIONS FOR AN INHOMOGENEOUS VON KÁRMÁN SPECTRUM

The derived Eqs. (14), (16), and (17) allow us to study the effects of turbulence inhomogeneity and anisotropy on the coherence function and mean sound field provided that the three-dimensional spectra of inhomogeneous anisotropic turbulence are known. The exact forms for such spectra have not yet been developed. Therefore, to study these effects, one needs to use model spectra of inhomogeneous and/or anisotropic turbulence.

Mann's spectrum[22] provides a model spectrum of anisotropic shear driven turbulence. Using this spectrum and Eq. (14), the effects of turbulence anisotropy on the coherence function of a plane sound wave were studied numerically in Refs. 18 and 19. The results obtained clearly show that turbulence anisotropy significantly affects the coherence of a plane wave.

In this section, we will use a model inhomogeneous isotropic von Kármán spectrum of velocity fluctuations to study the effects of inhomogeneity of atmospheric turbulence on the mean sound field and coherence functions of plane and spherical sound waves. Hereinafter, we ignore temperature fluctuations. Von Kármán's form of the energy spectrum of velocity fluctuations is given by[23,24]

$$E(K) = \frac{55\Gamma(\frac{5}{6})}{9\sqrt{\pi}\Gamma(\frac{1}{3})}\frac{\sigma^2 K^5 \ell^5}{(1+K^2\ell^2)^{17/6}}, \tag{18}$$

where Γ is the gamma function, σ^2 is the variance of a single velocity component, and ℓ is a length scale that is roughly the size of the largest eddies in the field. We refer to ℓ as the outer length scale in the following discussion. For shear- and buoyancy-driven turbulence, the von Kármán spectrum, Eq. (18) allows one to realistically capture the dependence of σ^2 and ℓ on height.[23,24] This spectrum, however, does not account for anisotropy of turbulence.

Let us calculate the effective three-dimensional spectral density Φ_{eff} appearing in Eqs. (14), (16), and (17) for the von Kármán spectrum. The velocity spectrum $\Phi_{ij}(\mathbf{K})$ is expressed in terms of the energy spectrum E by a well-known formula:

$$\Phi_{ij}(\mathbf{K}) = \frac{E(K)}{4\pi K^4}(K^2\delta_{ij} - K_i K_j). \tag{19}$$

Using this formula and noting that $\Phi_{11} = \Phi_{xx}$, we obtain the following expression for the three-dimensional spectral density

$$\Phi_{\text{eff}}(\eta x,0;0,K_\perp) = \frac{4\Phi_{11}(\eta x,0;0,K_\perp)}{c_0^2} = \frac{55\Gamma(\frac{5}{6})}{9\pi^{3/2}\Gamma(\frac{1}{3})c_0^2}\frac{\sigma^2(\eta)K_\perp^2\ell^5(\eta)}{(1+K_\perp^2\ell^2(\eta))^{17/6}}. \tag{20}$$

With this effective three-dimensional spectral density, we derive the following result for the mean sound field by performing the integration over $\mathbf{K}_\perp$ in Eq. (17):

$$\langle p(x,\mathbf{r})\rangle = p_0(x,\mathbf{r}) \times \exp\left[-\frac{\sqrt{\pi}\Gamma(\frac{5}{6})k^2 x}{\Gamma(\frac{1}{3})c_0^2} \int_0^1 d\eta\, \sigma^2(\eta)\ell(\eta) \right]. \quad (21)$$

The plane-wave coherence function, from Eq. (14), is

$$\Gamma^p(x;\mathbf{r}_+,\mathbf{r}_-) = |A_0|^2 \exp\left[-\frac{2\sqrt{\pi}\Gamma(\frac{5}{6})k^2 x}{\Gamma(\frac{1}{3})c_0^2} \times \int_0^1 d\eta\, \sigma^2(\eta)\ell(\eta) \times \left\{ 1 - \frac{\Gamma(\frac{1}{6})}{\pi}\left[\frac{r_-}{2\ell(\eta)}\right]^{5/6} K_{5/6}\left[\frac{r_-}{\ell(\eta)}\right] - \left[\frac{r_-}{2\ell(\eta)}\right] K_{1/6}\left[\frac{r_-}{\ell(\eta)}\right] \right\} \right], \quad (22)$$

where $K_{5/6}$ and $K_{1/6}$ are the modified Bessel functions. The spherical-wave coherence function, from Eq. (16), is

$$\Gamma^s(x;\mathbf{r}_+,\mathbf{r}_-) = \frac{1}{(4\pi x)^2} \exp\left[\frac{ik\mathbf{r}_+\cdot\mathbf{r}_-}{x} - \frac{2\sqrt{\pi}\Gamma(\frac{5}{6})k^2 x}{\Gamma(\frac{1}{3})c_0^2} \times \int_0^1 d\eta\, \sigma^2(\eta)\ell(\eta) \times \left\{ 1 - \frac{\Gamma(1/6)}{\pi}\left[\frac{\eta r_-}{2\ell(\eta)}\right]^{5/6} K_{5/6}\left[\frac{\eta r_-}{\ell(\eta)}\right] - \left[\frac{\eta r_-}{2\ell(\eta)}\right] K_{1/6}\left[\frac{\eta r_-}{\ell(\eta)}\right] \right\} \right]. \quad (23)$$

When σ^2 and ℓ are constant, these expressions reduce to previous ones for propagation through turbulent velocity fluctuations described by a homogeneous von Kármán spectrum.[13,4]

The height dependence of variances and length scales of atmospheric turbulence, in the context of sound propagation modeling with von Kármán's spectrum, is discussed in detail in Refs. 23 and 24. A case of particular practical importance is propagation through turbulence produced by a ground-based shear in the atmospheric surface layer. For this situation, the variance is nearly independent of the height h from the ground, whereas the outer length scale is described by $\ell = bh$, where $b \simeq 1.8$ is an empirical constant. Let us consider propagation along a straight line through a layer with these properties. Since the outer length scale is proportional to height, it varies linearly along the straight line path from its value $\ell_1 = bh_1$ at the start to its value $\ell_2 = bh_2$ at the end; that is, $\ell(\eta) = (1-\eta)\ell_1 + \eta\ell_2$. If we define θ to be the elevation angle of the path and $\ell_m = (\ell_1 + \ell_2)/2$ as the length scale at the midpoint, we have $\sin\theta = (h_2 - h_1)/x$ and, therefore,

$$\ell(\eta) = \ell_m + b(\eta - 1/2)x \sin\theta. \quad (24)$$

The mean field is readily determined by substituting Eq. (24) into (21) and integrating, with the result

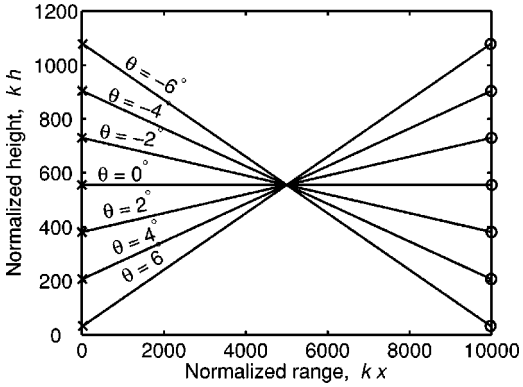


FIG. 2. Transmission paths intersecting at a common midpoint, as used for the calculations in Figs. 3 and 4. Positive elevation angles θ are for the receiving array higher than the source. All dimensions are shown normalized by the acoustic wavenumber. The height of the paths at the midpoint is $kh = 556$. Sources are indicated by an "×" and centers of receiving arrays are indicated by circles. The scale of the vertical axis is exaggerated for illustrative purposes.

$$\langle p(x,\mathbf{r})\rangle = p_0(x,\mathbf{r}) \exp\left[-\frac{k^2 x \sigma^2 \ell_m}{c_0^2 B(\frac{1}{2},\frac{1}{3})} \right], \quad (25)$$

where B is the beta function. Note that the mean field is independent of the elevation angle θ.

In contrast, the integrals in Eqs. (22) and (23) for the plane-wave and spherical-wave coherence cannot be solved in analytical form. Furthermore, given that $\ell(\eta)$ is an argument of the Bessel functions, one would not expect the results to be independent of the elevation angle θ. Although the coherence does depend on θ, it can be shown from Eq. (22) that the plane-wave coherence is reciprocal; that is, the coherence for $\theta = \alpha$ is the same as for $\theta = -\alpha$, so long as the values of ℓ_m and x for the two paths match. (The proof simply involves changing variables from η to $\eta' = 1 - \eta$ in the integral for the case $\theta = -\alpha$, which makes the integral identical to the case $\theta = \alpha$.) The spherical-wave coherence, however, is not reciprocal. Symmetry is broken by the factor ηr_- under the integral, which weights turbulence properties along the sound propagation path in different manners. Such weighting is well know in theories of wave propagation through homogeneous isotropic turbulence, e.g., see Ref. 6. However, nonreciprocity in the spherical-wave coherence can only occur for the case of inhomogeneous turbulence.

Example calculations of Γ^p and Γ^s, along the family of propagation paths illustrated in Fig. 2, are shown in Figs. 3 and 4. Note that in these calculations, all of the propagation paths have the same length and intersect at the midway point ($\eta = 0.5$). For these calculations, we have set $kx = 10^5$, $k\ell_m = 10^3$, $\sigma^2/c_0^2 = 10^{-5}$, and $b = 1.8$ in Eqs. (22)–(24). The coherence equations were integrated numerically by applying the trapezoidal rule along 100 equally spaced points between $\eta = 0$ and $\eta = 1$. Curves corresponding to $kr_- = 10, 20, \ldots, 100$ are shown. As expected, the calculations for Γ^p (Fig. 3) do not depend on the sign of θ. The paths with steeper elevation angles have somewhat lower coherence than the nearly horizontal ones. The primary reason is that the steeper paths pass through regions having larger outer length scales, which decreases the coherence. The calculations for Γ^s (Fig. 4) ex-

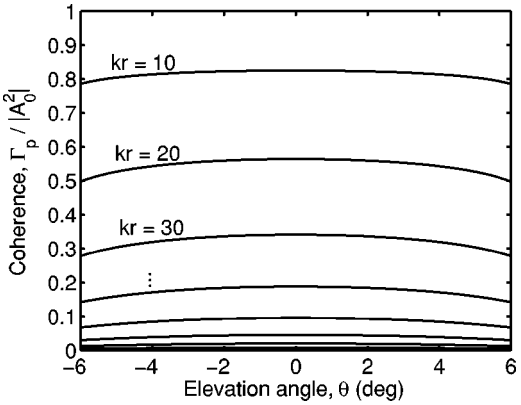


FIG. 3. Normalized plane-wave coherence as a function of elevation angle. The transmission paths intersect at a common midpoint as illustrated in Fig. 2. Curves are shown for sensor separations $kr_- = 10,20,...,100$.

hibit an interesting asymmetry with respect to θ: Negatively inclined propagation paths (source higher than the receiving array) exhibit lower coherence. The change in coherence can be quite substantial. At $kr_- = 60$, for example, coherence for the $\theta = 6°$ path is 16% larger than for the $\theta = 0°$ path, whereas coherence for the $\theta = -6°$ path is 41% smaller than for the $\theta = 0°$ path. If, in this example, the sound frequency is 3 kHz, then the propagation distance x would be 1.8 km, the scale ℓ_m would be 18 m, and the separation between receivers r_- would be 1 m.

An additional set of calculations of Γ^p and Γ^s, along the family of slanted paths illustrated in Fig. 5, are shown in Figs. 6 and 7. For these paths, the source is fixed at a height near the ground such that $k\ell_1 = 10$ for $\theta > 0$, and the receiver is fixed at $k\ell_2 = 10$ for $\theta < 0$. Therefore, $\theta = 0$ is essentially a ground-to-ground transmission, $\theta > 0$ is a ground-to-air transmission, and $\theta < 0$ is an air-to-ground transmission. The length scale for the ground-based source has the following dependence on η:

$$\ell(\eta) = \ell_1 + b\,\eta x \sin\theta. \tag{26}$$

Similarly, for the ground-based receiver, we set

$$\ell(\eta) = \ell_2 + b\,\eta x \sin\theta. \tag{27}$$

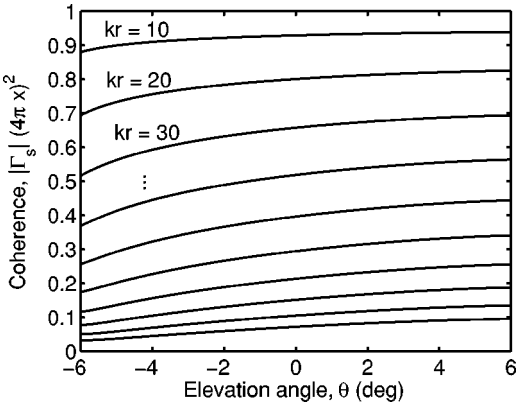


FIG. 4. Normalized spherical-wave coherence as a function of elevation angle. The transmission paths intersect at a common midpoint as illustrated in Fig. 2. Curves are shown for sensor separations $kr_- = 10,20,...,100$.

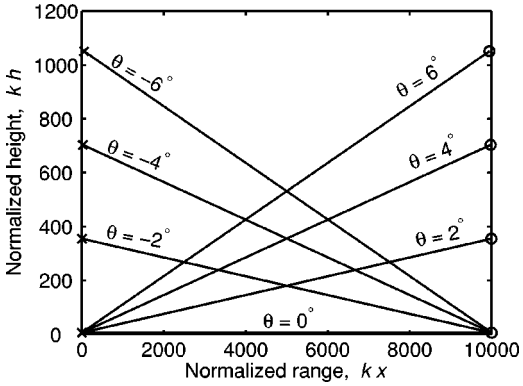


FIG. 5. Transmission paths starting or ending at the ground, as used for the calculations in Figs. 6 and 7. Positive elevation angles θ are for the receiving array higher than the source. All dimensions are shown normalized by the acoustic wavenumber. Sources are indicated by an "×" and centers of receiving arrays are indicated by circles. The scale of the vertical axis is exaggerated for illustrative purposes.

One prominent feature evident in Figs. 6 and 7 is the presence of slant angles for which the coherence is minimized. Again, a pronounced nonreciprocity is observed in the spherical-wave coherence. For both the plane and spherical waves, coherence minima are observed for horizontal (ground-to-ground) transmission when the sensor separation is comparable to or smaller than the length scale ℓ_1. As the separation is increased beyond ℓ_1, the coherence minima shift to slant angles slightly off the horizontal. The cause of this behavior is, first, that the outer length scale of the turbulence is very small for ground-to-ground transmission. As a result, even for sensor separations larger than the outer length scale, the coherence remains relatively high. As the transmission path becomes more slanted, propagation occurs in areas of the turbulence possessing larger outer length scales, which causes the coherence to diminish. Eventually, for very slanted paths, a significant portion of the propagation occurs in regions where the sensor separation is small compared to the outer length scale, which has the effect of improving coherence.

VI. CONCLUSIONS

Line-of-sight sound propagation through inhomogeneous anisotropic turbulence with temperature and velocity

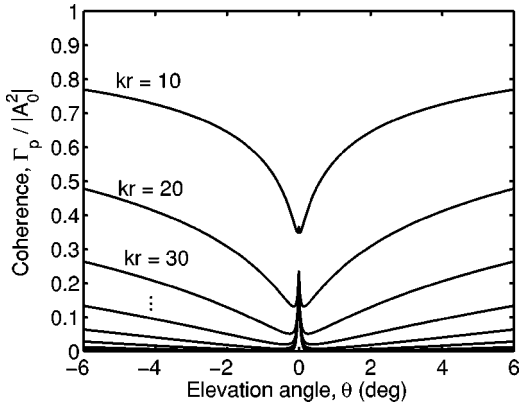


FIG. 6. Normalized plane-wave coherence as a function of elevation angle. The transmission paths start or end at the ground as illustrated in Fig. 5. Curves are shown for sensor separations $kr_- = 10,20,...,100$.

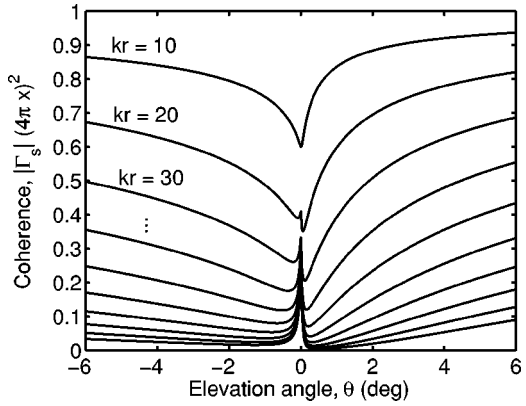


FIG. 7. Normalized spherical-wave coherence as a function of elevation angle. The transmission paths start or end at the ground as illustrated in Fig. 5. Curves are shown for sensor separations $kr_- = 10,20,...,100$.

fluctuations has been studied. Starting from the equations for the first two statistical moments of a sound field obtained in our earlier paper,[20] we derived equations for the coherence function and the mean sound field of initially arbitrary waveform propagating through inhomogeneous anisotropic turbulence, Eqs. (12) and (17), respectively. Using the former equation, we obtained formulas for the coherence functions of plane and spherical sound waves, Eqs. (14) and (16). All these equations are new results obtained in the paper. These equations coincide with those known in the literature in two limiting cases: (a) homogeneous isotropic turbulence with temperature and velocity fluctuations, and (b) inhomogeneous anisotropic turbulence with temperature fluctuations only.

Using the inhomogeneous von Kármán spectrum of velocity fluctuations, we numerically studied the coherence functions of plane and spherical sound waves propagating in a turbulent atmosphere. We showed that the coherence functions significantly depend on the turbulence inhomogeneity. In particular, we revealed nonreciprocity for spherical wave propagation: Coherence is generally better when the source is near the ground and the receiver is elevated, rather than vice versa.

ACKNOWLEDGMENTS

This material is partly based upon work supported by the U.S. Army Research Office under Contract No. DAAD19-01-1-0640.

APPENDIX A: APPROXIMATE SOLUTION OF EQ. (7)

In this Appendix, we present an approximate solution of Eq. (7). This solution is similar (but not identical) to that developed in Ref. 10. The difference between these solutions is that we consider the case of inhomogeneous turbulence with temperature and velocity fluctuations, while in Ref. 10 the case of homogeneous turbulence with temperature fluctuations only was discussed.

In Eq. (7), we express the functions $\Gamma(x;\mathbf{r}_+,\mathbf{r}_-)$ and $M(x;\mathbf{r}_+,\mathbf{r}_-)$ in the form of the Fourier integrals with respect to the argument $\mathbf{r}_+$:

$$\Gamma(x;\mathbf{r}_+,\mathbf{r}_-)=\int d\mathbf{q}e^{i\mathbf{q}\cdot\mathbf{r}_+}g(x;\mathbf{q},\mathbf{r}_-), \tag{A1}$$

and

$$M(x;\mathbf{r}_+,\mathbf{r}_-)=\int d\mathbf{q}e^{-i\mathbf{q}\cdot\mathbf{r}_+}m(x;\mathbf{q},\mathbf{r}_-). \tag{A2}$$

Here, $g(x;\mathbf{q},\mathbf{r}_-)$ and $m(x;\mathbf{q},\mathbf{r}_-)$ are the corresponding spectral densities. Note that, in the integrand of Eq. (A2), the minus sign in the exponent is chosen for convenience of subsequent calculations.

Substitution of Eqs. (A1) and (A2) into Eq. (7) results in the following integro-differential equation for the spectral density $g(x;\mathbf{q},\mathbf{r}_-)$:

$$\frac{\partial g(x;\mathbf{q},\mathbf{r}_-)}{\partial x}+\frac{\mathbf{q}}{k}\cdot\frac{\partial g(x;\mathbf{q},\mathbf{r}_-)}{\partial \mathbf{r}_-}+\frac{k^2}{8}\int d\mathbf{q}_1 g(x;\mathbf{q}_1+\mathbf{q},\mathbf{r}_-)m(x;\mathbf{q}_1,\mathbf{r}_-)=0. \tag{A3}$$

This equation can be written in the equivalent form:

$$\frac{\partial}{\partial x}\left[\exp\left(\frac{\mathbf{q}x}{k}\cdot\frac{\partial}{\partial \mathbf{r}_-}\right)g(x;\mathbf{q},\mathbf{r}_-)\right]=-\frac{k^2}{8}\exp\left(\frac{\mathbf{q}x}{k}\cdot\frac{\partial}{\partial \mathbf{r}_-}\right)\times\int d\mathbf{q}_1 g(x;\mathbf{q}_1+\mathbf{q},\mathbf{r}_-)m(x;\mathbf{q}_1,\mathbf{r}_-). \tag{A4}$$

Indeed, by operating with $\partial/\partial x$ on both terms on the left-hand side of Eq. (A4), we obtain Eq. (A3). In Eq. (A4), the pseudo-differential operator $\exp(\mathbf{q}x/k\cdot\partial/\partial\mathbf{r}_-)$ has the following property:

$$\exp\left(\frac{\mathbf{q}x}{k}\cdot\frac{\partial}{\partial \mathbf{r}_-}\right)F(\mathbf{r}_-)=F\left(\mathbf{r}_-+\frac{\mathbf{q}x}{k}\right), \tag{A5}$$

where $F(\mathbf{r}_-)$ is an arbitrary function of $\mathbf{r}_-$. Using Eq. (A5), integro-differential equation (A4) can be written as

$$\frac{\partial g(x;\mathbf{q},\mathbf{r}_-+\mathbf{q}x/k)}{\partial x}=-\frac{k^2}{8}\int d\mathbf{q}_1 g(x;\mathbf{q}_1+\mathbf{q},\mathbf{r}_-+\mathbf{q}x/k)m(x;\mathbf{q}_1,\mathbf{r}_-+\mathbf{q}x/k). \tag{A6}$$

Equation (A6) is equivalent to Eq. (7). Both equations cannot be solved analytically in a general case of inhomogeneous turbulence, when the function M depends on the argument $\mathbf{r}_+$.

Let us consider the spectral density $m(x;\mathbf{q},\mathbf{r}_-)$ appearing in Eq. (A6). Using Eq. (A2), $m(x;\mathbf{q},\mathbf{r}_-)$ is expressed in terms of the function $M(x;\mathbf{r}_+,\mathbf{r}_-)$. Then, M is replaced by the right-hand side of Eq. (11). As a result, we have

$$\begin{aligned} m(x;\mathbf{q},\mathbf{r}_-)&=\frac{1}{4\pi^2}\int d\mathbf{r}_+ e^{i\mathbf{q}\cdot\mathbf{r}_+}M(x;\mathbf{r}_+,\mathbf{r}_-)\\ &=\frac{1}{2\pi}\int d\mathbf{r}_+ e^{i\mathbf{q}\cdot\mathbf{r}_+}\int d\mathbf{K}_\perp[\Phi_{\mathrm{eff}}(x,\mathbf{r}_+ +\mathbf{r}_-/2;0,\mathbf{K}_\perp)+\mathbf{\Phi}_{\mathrm{eff}}(x,\mathbf{r}_+ -\mathbf{r}_-/2;0,\mathbf{K}_\perp)\\ &\quad -2e^{i\mathbf{K}_\perp\cdot\mathbf{r}_-}\mathbf{\Phi}_{\mathrm{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp)]. \end{aligned} \tag{A7}$$

Substituting the value of the function m into Eq. (A6), we have

$$\frac{\partial g(x;\mathbf{q},\mathbf{r}_- + \mathbf{q}x/k)}{\partial x} = -\frac{k^2}{16\pi}\int d\mathbf{q}_1 \int d\mathbf{K}_\perp \int d\mathbf{r}_+ g(x;\mathbf{q}_1 + \mathbf{q},\mathbf{r}_- + \mathbf{q}x/k)\times e^{i\mathbf{q}_1\cdot\mathbf{r}_+}[\Phi_{\text{eff}}(x,\mathbf{r}_+ + (\mathbf{r}_- + x\mathbf{q}/k)/2;0,\mathbf{K}_\perp) + \Phi_{\text{eff}}(x,\mathbf{r}_+ - (\mathbf{r}_- + x\mathbf{q}/k)/2;0,\mathbf{K}_\perp) - 2e^{i\mathbf{K}_\perp\cdot(\mathbf{r}_- + x\mathbf{q}/k)}\Phi_{\text{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp)]. \tag{A8}$$

Let us consider three integrals over $\mathbf{r}_+$ corresponding to three functions Φ_{eff} in the square brackets in Eq. (A8). The last of these integrals is given by

$$I = -2\int d\mathbf{r}_+ g(x;\mathbf{q}_1 + \mathbf{q},\mathbf{r}_- + \mathbf{q}x/k) \times e^{i\mathbf{q}_1\cdot\mathbf{r}_+} e^{i\mathbf{K}_\perp\cdot(\mathbf{r}_- + x\mathbf{q}/k)}\Phi_{\text{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp). \tag{A9}$$

It can be shown that, in Eqs. (A8) and (A9), the scale of $g(x;\mathbf{q}_1+\mathbf{q},\mathbf{r}_- + \mathbf{q}x/k)$ with respect to $\mathbf{q}_1$ is $\sqrt{k/x}$, which is the inverse of the first Fresnel zone. Let a be the scale of variations of $\Phi_{\text{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp)$ with respect to the argument $\mathbf{r}_+$. If $a \gg \sqrt{x/k}$, the factor $e^{i\mathbf{q}_1\cdot\mathbf{r}_+}$ in the integrand in Eq. (A9) oscillates rapidly in the region $r_+ > a$ so that a contribution from this region to I can be ignored. In the region $r_+ < a$, the argument $\mathbf{r}_+$ of $\Phi_{\text{eff}}(x,\mathbf{r}_+;0,\mathbf{K}_\perp)$ can be set to 0 and the integral over $\mathbf{r}_+$ in Eq. (A9) can be evaluated:

$$I = -8\pi^2\delta(\mathbf{q}_1) g(x;\mathbf{q}_1 + \mathbf{q},\mathbf{r}_- + \mathbf{q}x/k) \times e^{i\mathbf{K}_\perp\cdot(\mathbf{r}_- + x\mathbf{q}/k)}\Phi_{\text{eff}}(x,0;0,\mathbf{K}_\perp). \tag{A10}$$

The other two integrals over $\mathbf{r}_+$ on the right-hand side of Eq. (A8) can be evaluated similarly. As a result, Eq. (A8) takes the form

$$\frac{\partial g(x;\mathbf{q},\mathbf{r}_- + \mathbf{q}x/k)}{\partial x} = -\frac{k^2}{8} g(x;\mathbf{q},\mathbf{r}_- + \mathbf{q}x/k) \times \tilde{M}(x;0,\mathbf{r}_- + \mathbf{q}x/k), \tag{A11}$$

where

$$\tilde{M}(x;0,\mathbf{r}_-) = 4\pi\int d\mathbf{K}_\perp [1 - e^{i\mathbf{K}_\perp\cdot\mathbf{r}_-}]\Phi_{\text{eff}}(x,0;0,\mathbf{K}_\perp). \tag{A12}$$

Equation (A11) can be solved for g:

$$g(x;\mathbf{q},\mathbf{r}_- + \mathbf{q}x/k) = g_0(\mathbf{q},\mathbf{r}_-)\exp\left[-\frac{k^2}{8}\int_0^x dx' \tilde{M}(x';0,\mathbf{r}_- + \mathbf{q}x'/k)\right]. \tag{A13}$$

Here, the function $g_0(\mathbf{q},\mathbf{r}_-) = g(x=0,\mathbf{q},\mathbf{r}_-)$. This function is related to the function Γ_0 by a formula similar to Eq. (A1):

$$\Gamma_0(\mathbf{r}_+,\mathbf{r}_-) = \int d\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{r}_+} g_0(\mathbf{q},\mathbf{r}_-). \tag{A14}$$

In Eq. (A13), we introduce a new variable: $\mathbf{r}'_- = \mathbf{r}_- + \mathbf{q}x/k$:

$$g(x;\mathbf{q},\mathbf{r}'_-) = g_0(\mathbf{q},\mathbf{r}'_- - \mathbf{q}x/k)\exp\left[-\frac{k^2}{8}\int_0^x dx' \times \tilde{M}(x';0,\mathbf{r}'_- + \mathbf{q}(x'-x)/k)\right]. \tag{A15}$$

In this equation, we change from $\mathbf{r}'_-$ to $\mathbf{r}_-$. Furthermore, using Eq. (A14), we express g_0 in terms of Γ_0. Finally, substituting the right-hand side of the resulting equation into Eq. (A1), we obtain a formula for the coherence function:

$$\Gamma(x;\mathbf{r}_+,\mathbf{r}_-) = \frac{1}{4\pi^2}\int d\mathbf{q}\int d\mathbf{r}'_+ \Gamma_0(\mathbf{r}'_+,\mathbf{r}_- - \mathbf{q}x/k) \times \exp\left[i\mathbf{q}\cdot(\mathbf{r}_+ - \mathbf{r}'_+) - \frac{k^2}{8} \times \int_0^x dx' \tilde{M}(x';0,\mathbf{r}_- - \mathbf{q}(x-x')/k)\right]. \tag{A16}$$

In this formula, the integration variable $\mathbf{q}$ is replaced by a new variable $\mathbf{r}'_- = \mathbf{r}_- - \mathbf{q}x/k$. With this substitution, Eq. (A16) becomes Eq. (12).

APPENDIX B: DERIVATION OF A FORMULA FOR THE MEAN SOUND FIELD

In Appendix B, we present a derivation of a formula for the mean sound field.

It is convenient to express the mean sound field in the following form:

$$\langle p(x,\mathbf{r})\rangle = e^{ikx}\overline{A}(x,\mathbf{r}). \tag{B1}$$

Here, $\overline{A}$ is the mean value of the complex amplitude of the sound field.

A closed equation for $\overline{A}$ was derived in Ref. 20 for the case of sound propagation in a refractive turbulent atmosphere near an impedance boundary. For the considered case of line-of-sight sound propagation, this equation simplifies and has the form

$$\frac{\partial\overline{A}(x,\mathbf{r})}{\partial x} - \frac{i}{2k}\frac{\partial^2\overline{A}(x,\mathbf{r})}{\partial\mathbf{r}^2} + \frac{k^2}{8} b_{\text{eff}}(x;\mathbf{r},\mathbf{r})\overline{A}(x,\mathbf{r}) = 0. \tag{B2}$$

In Eq. (B2), $b_{\text{eff}}(x;\mathbf{r},\mathbf{r})$ is given by [see Eq. (9)]

$$b_{\text{eff}}(x;\mathbf{r},\mathbf{r}) = 2\pi\int d\mathbf{K}_\perp \Phi_{\text{eff}}(x,\mathbf{r};0,\mathbf{K}_\perp). \tag{B3}$$

An initial condition to Eq. (B2) is formulated in the plane $x=0$:

$$\overline{A}(x=0,\mathbf{r}) = p_0(\mathbf{r}). \tag{B4}$$

Let us express the functions $\overline{A}(x,\mathbf{r})$ and $b_{\text{eff}}(x;\mathbf{r},\mathbf{r})$ as the Fourier integrals:

$$\overline{A}(x,\mathbf{r})=\int d\mathbf{q}e^{i\mathbf{q}\cdot\mathbf{r}}a(x,\mathbf{q}), \tag{B5}$$

$$b_{\text{eff}}(x;\mathbf{r},\mathbf{r})=\int d\mathbf{q}e^{-i\mathbf{q}\cdot\mathbf{r}}b(x;\mathbf{q}), \tag{B6}$$

where $a(x;\mathbf{q})$ and $b(x;\mathbf{q})$ are the corresponding spectral densities. Substituting Eqs. (B5) and (B6) into Eq. (B2), we obtain an integro-differential equation for the spectral density $a(x;\mathbf{q})$:

$$\frac{\partial a(x,\mathbf{q})}{\partial x}+\frac{iq^2a(x,\mathbf{q})}{2k}+\frac{k^2}{8}\int d\mathbf{q}_1a(x,\mathbf{q}+\mathbf{q}_1)b(x,\mathbf{q}_1)=0. \tag{B7}$$

Let us consider the spectral density $b(x,\mathbf{q})$ appearing in Eq. (B7). Using Eq. (B6), we express $b(x,\mathbf{q})$ in terms of $b_{\text{eff}}(x,\mathbf{q})$. Then, we replace $b_{\text{eff}}(x,\mathbf{q})$ by the right-hand side of Eq. (B3). As a result, Eq. (B7) takes the form

$$\frac{\partial a(x,\mathbf{q})}{\partial x}+\frac{iq^2a(x,\mathbf{q})}{2k}+\frac{k^2}{16\pi}\times\int d\mathbf{q}_1\int d\mathbf{K}_\perp\int d\mathbf{r}e^{i\mathbf{q}_1\cdot\mathbf{r}}a(x,\mathbf{q}+\mathbf{q}_1)\times\Phi_{\text{eff}}(x,\mathbf{r};0,\mathbf{K}_\perp)=0. \tag{B8}$$

The integral over $\mathbf{r}$ in the integrand of this equation can be evaluated similarly to that in Eq. (A9). The result is

$$\int d\mathbf{r}e^{i\mathbf{q}_1\cdot\mathbf{r}}a(x,\mathbf{q}+\mathbf{q}_1)\Phi_{\text{eff}}(x,\mathbf{r};0,\mathbf{K}_\perp)=4\pi^2\delta(\mathbf{q}_1)a(x,\mathbf{q}+\mathbf{q}_1)\Phi_{\text{eff}}(x,0;0,\mathbf{K}_\perp). \tag{B9}$$

Using this result in Eq. (B8), we have

$$\frac{\partial a(x,\mathbf{q})}{\partial x}+\frac{iq^2a(x,\mathbf{q})}{2k}+\frac{k^2}{8}a(x,\mathbf{q})b_{\text{eff}}(x;0,0)=0. \tag{B10}$$

Here,

$$b_{\text{eff}}(x;0,0)=2\pi\int d\mathbf{K}_\perp\Phi_{\text{eff}}(x,0;0,\mathbf{K}_\perp). \tag{B11}$$

Equation (B10) can be solved for a:

$$a(x,\mathbf{q})=a_0(\mathbf{q})\exp\left[-\frac{iq^2x}{2k}-\frac{k^2}{8}\int_0^x dx'b_{\text{eff}}(x';0,0)\right]. \tag{B12}$$

Here, the function $a_0(\mathbf{q})=a(x=0,\mathbf{q})$ is the Fourier transform of $p_0(\mathbf{r})$:

$$p_0(\mathbf{r})=\int d\mathbf{q}e^{i\mathbf{q}\cdot\mathbf{r}}a_0(\mathbf{q}). \tag{B13}$$

Substituting the spectral density a given by Eq. (B12) into Eq. (B5) and replacing $b_{\text{eff}}(x';0,0)$ by its value from Eq. (B11), we get

$$\overline{A}(x,\mathbf{r})=\exp\left[-\frac{\pi k^2}{4}\int_0^x dx'\int d\mathbf{K}_\perp\Phi_{\text{eff}}(x',0;0,\mathbf{K}_\perp)\right]\times\int d\mathbf{q}a_0(\mathbf{q})\exp(i\mathbf{q}\cdot\mathbf{r}-iq^2x/2k). \tag{B14}$$

In a medium without random inhomogeneities, $\Phi_{\text{eff}}(x',0;0,\mathbf{K}_\perp)=0$. Therefore, the integral on the right-hand side of Eq. (B14) is equal to the complex amplitude $A_0(x,\mathbf{r})$ of a sound field in a medium without random inhomogeneities. Thus, Eq. (B14) can be written in the following form:

$$\overline{A}(x,\mathbf{r})=\exp\left[-\frac{\pi k^2}{4}\int_0^x dx'\int d\mathbf{K}_\perp\Phi_{\text{eff}}(x',0;0,\mathbf{K}_\perp)\right]\times A_0(x,\mathbf{r}). \tag{B15}$$

A formula for the mean sound field, Eq. (17), can be obtained by eliminating $\overline{A}$ between Eqs. (B1) and (B15).

[1] E. M. Salomons, *Computational Atmospheric Acoustics* (Kluwer Academic, Dordrecht, 2001).
[2] V. E. Ostashev, P. Blanc-Benon, and D. Juvé, "Coherence function of a spherical acoustic wave after passing through a turbulent jet," C. R. Acad. Sci., Ser. IIb: Mec., Phys., Chim., Astron. **326**(1), 39–45 (1998).
[3] S. L. Collier and D. K. Wilson, "Performance bounds for passive sensor arrays operating in a turbulent medium: Plane wave analysis," J. Acoust. Soc. Am. **113**, 2704–2718 (2001).
[4] D. K. Wilson, "Performance bounds for acoustic angle-of-arrival arrays operating in atmospheric turbulence," J. Acoust. Soc. Am. **103**, 1306–1319 (1998).
[5] V. Mellert and B. Schwarz-Rohr, "Correlation and coherence measurements of a spherical wave traveling in the atmospheric boundary layer," in *Proceedings 7th Int. Symp. Long Range Sound Propagation* (Ecole Centrale de Lyon, Lyon, France, 1996), pp. 391–405.
[6] V. I. Tatarskii, *The Effects of the Turbulent Atmosphere on Wave Propagation* (Keter, Jerusalem, 1971).
[7] L. A. Chernov, *Waves in Randomly-inhomogeneous Media* (in Russian) (Nauka, Moscow, 1975).
[8] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978).
[9] S. M. Flatte, R. Dashen, W. H. Munk, K. M. Watson, and F. Zachariasen, *Sound Transmission through a Fluctuating Ocean* (Cambridge U.P., New York, 1979).
[10] S. M. Rytov, Y. A. Kravtsov, and V. I. Tatarskii, *Principles of Statistical Radio Physics. Part 4, Wave Propagation through Random Media* (Springer, Berlin, 1989).
[11] H. Sato and M. C. Fehler, *Seismic Wave Propagation and Scattering in the Heterogeneous Earth* (Springer, New York, 1997).
[12] A. D. Wheelon, *Electromagnetic Scintilation. I. Geometrical Optics* (Cambridge U.P., Cambridge, 2001).
[13] V. E. Ostashev, *Acoustics in Moving Inhomogeneous Media* (E&FN SPON, London, 1997).
[14] M. Kelly, R. Raspet, and H. E. Bass, "Scattering of sonic booms by anisotropic turbulence in the atmosphere," J. Acoust. Soc. Am. **107**, 3059–3064 (2000).
[15] V. E. Ostashev and D. K. Wilson, "Log-amplitude and phase fluctuations of a plane wave propagating through anisotropic, inhomogeneous turbulence," Acust. Acta Acust. **87**, 685–694 (2001).
[16] V. E. Ostashev and D. K. Wilson, "Spherical wave propagation through inhomogeneous, anisotropic turbulence: log-amplitude and phase correlations," J. Acoust. Soc. Am. (in print).
[17] D. K. Wilson, "Calculated coherence and extinction of sound waves propagating through anisotropic, shear-induced turbulent velocity fluctuations," J. Acoust. Soc. Am. **105**, 658–671 (1999).
[18] V. E. Ostashev and D. K. Wilson, "Line-of-sight sound propagation through anisotropic and inhomogeneous atmospheric turbulence," in *Proceedings of Int. Workshop: Tomograpy and Acoustics: Recent Developments and Methods*, Leipzig, Germany, 2001, pp. 63–70.
[19] V. E. Ostashev and D. K. Wilson, "Sound propagation through inhomogeneous, anisotropic, and intermittent atmospheric turbulence," in *Proc.*

10th Int. Symposium on Long Range Sound Propagation, Grenoble, France, 2002, pp. 236–248.
[20] D. K. Wilson and V. E. Ostashev, ''Statistical moments of the sound field propagating in a random, refractive medium near an impedance boundary,'' J. Acoust. Soc. Am. **109**, 1909–1922 (2001).
[21] A. S. Gurvich, A. I. Kon, V. L. Mironov, and S. S. Khmelevtsov, *Laser Radiation in the Turbulent Atmosphere* (Nauka, Moscow, 1976) (in Russian).
[22] J. Mann, ''The spatial structure of neutral atmospheric surface layer turbulence,'' J. Fluid Mech. **273**, 141–168 (1994).
[23] D. K. Wilson, ''A turbulence spectral model for sound propagation in the atmosphere that incorporates shear and buoyancy forcings,'' J. Acoust. Soc. Am. **108**, 2021–2038 (2000).
[24] G. H. Goedecke, V. E. Ostashev, D. K. Wilson, and H. J. Auvermann, ''Quasi-wavelet model of von Karman spectrum of turbulent velocity fluctuations,'' Boundary Layer Meteorology (in print).

Experimental study of the Doppler shift generated by a vibrating scatterer

Régis Wunenburger,[a] Nicolás Mujica,[b] and Stéphan Fauve
Laboratoire de Physique Statistique, Ecole Normale Supérieure, CNRS UMR 8550, 24 rue Lhomond, 75231 Paris Cedex 05, France

(Received 9 May 2003; revised 28 October 2003; accepted 31 October 2003)

We report an experimental study of the backscattering of a sound wave of frequency f by a surface vibrating harmonically at frequency F ($F \ll f$) and amplitude A in the regime where the Doppler effect overcomes bulk nonlinear effects. When the duration t_0 of the analyzed time series of the scattered wave is small compared to the vibration period, the power spectrum of the backscattered wave is proportional to the probability density function of the scatterer velocity, which presents two peaks shifted from f by roughly $2fA\Omega/c$ ($\Omega = 2\pi F$). On the contrary, when $t_0 \gg F^{-1}$, sidebands at frequencies $f \pm nF$ (n integer) appear in the power spectrum, which are due to the phase modulation of the backscattered wave induced by its reflection on a moving boundary. We use the backscattered power spectrum to validate the phase modulation theory of the Doppler effect in the latter case for $2kA \ll 1$ and $2kA \gtrsim 1$ ($k = 2\pi f/c$, where c is the wave velocity) and we test the validity of an acoustic nonintrusive estimator of A as a function of power spectrum bandwidth and of A itself.
 [DOI: 10.1121/1.1635414]

PACS numbers: 43.20.Fn, 43.28.Py, 43.25.Lj [MFH] Pages: 507–514

I. INTRODUCTION

Velocity measurement using the Doppler shift of a wave reflected from a moving object is a widely used technique both with electromagnetic and acoustic waves. For objects moving at constant velocity, it is well known that the Doppler shift, calculated by means of coordinate transformation, is proportional to the velocity. The problem is more difficult when the motion is time dependent. In the case of a periodically oscillating object the problem was first carefully studied for electromagnetic waves.[1] The spectrum of the scattered wave is also modified due to the Doppler effect, which can be understood either as a nonlinear boundary condition imposed by the moving object[1,2] or as caused from the inhomogeneity in time of the moving medium that supports the wave propagation.[3] More precisely, the spectrum of the wave at frequency f scattered by a sinusoidally oscillating surface at frequency F is similar to that of a phase modulation process,[1] i.e., sidebands at frequencies $f \pm nF$ (n integer) appear in the spectrum of the scattered wave.

In the case of acoustic scattering, the situation is more complex, as the oscillating scatterer also emits a sound wave at frequency F which interacts with the scattered wave due to the nonlinear character of the equations of sound propagation.[4,5] Rogers remarked[6] that the bulk nonlinear wave mixing produces the same sideband peaks in the spectrum of the backscattered sound wave as the Doppler effect. Due to the lack of any decisive experiment, the criteria allowing us to discriminate between the two effects have been the subject of an intense debate,[3,6–11] and it has even been claimed that the contribution of the Doppler effect might be undetectable.[6,8–11] In a previous letter,[12] we showed that there exists a wide parameter range in which the Doppler shift gives the dominant contribution to the spectrum of the scattered wave and studied the cross-over to the bulk dominated nonlinear regime.

In this article we focus on the experimental situation where the Doppler effect induced by a vibrating scatterer overcomes bulk nonlinearities. In particular we study the backscattering of a high frequency f sound wave by a plane scatterer oscillating at low frequency $F \ll f$. We study the characteristic features of the Doppler effect in both the static and quasi-static regimes and we validate the phase modulation theory of the Doppler effect in the quasistatic regime for $2kA \ll 1$ and $2kA \gtrsim 1$. Many studies have been devoted to vibration measurements using ultrasonic techniques.[13–18] In particular, Huang *et al.* proposed an acoustic nonintrusive estimator of the scatterer oscillation amplitude based on the phase modulation of the backscattered wave.[15] To our knowledge, this estimator has not been validated experimentally. By comparing the scatterer oscillation amplitude obtained with both the acoustic nonintrusive estimator and vibration measurements performed with an accelerometer, we verify the accuracy of Huang *et al.*'s amplitude estimator.

This paper is organized as follows: in Sec. II we present the conditions for observing the static and quasi-static Doppler effect, and we recall some predictions concerning both surface Doppler and bulk nonlinear effects. This helps us to justify our choice of the experimental configuration, which is presented in Sec. III. In Sec. IV we present the main features of the static and quasi-static Doppler effect. In Sec. V we finally test the validity of a nonintrusive estimator of the

[a]Permanent address: Center de Physique Moléculaire Optique et Hertzienne, Université Bordeaux I, 351 cours de la Libération, 33405 Talence Cedex, France. Electronic mail: r.wunenburger@cpmoh.u-bordeaux1.fr
[b]Permanent address: Departamento de Física, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Av. Blanco Encalada 2008, Santiago, Chile. Present address: Nonlinear Dynamics Laboratory, Institute for Research in Electronics and Applied Physics, Bldg 223, University of Maryland, College Park, MD 20742.

scatterer oscillation amplitude. Conclusions are given in Sec. VI.

II. THEORETICAL CONSIDERATIONS

A. Static versus quasi-static Doppler effect

The usual picture of the Doppler effect is the constant frequency shift of an incident or emitted wave by an object moving at constant velocity V, which we will call the static Doppler effect. In the case of backscattering, the frequency shift Δf encountered by the scattered wave of frequency f is $2fV/c$. From an experimental point of view, when the velocity of the scatterer varies, the latter approach remains valid when the timescale of velocity variations is much larger than the duration t_0 of the scattered wave time series which is analyzed. For a periodic motion of frequency F, this implies $t_0 \ll F^{-1}$. In this case, the statistical distribution of the successively measured Doppler shifts is proportional to the probability density function (PDF) of the object velocity. In practice, since f and F are not commensurate, the computed power spectrum of the backscattered wave is the average of power spectra of many successive signal time series of duration t_0 ($t_0 \gg f^{-1}$) measured at random phase of the scatterer motion. Thus, this time-averaged power spectrum is expected to be proportional to the PDF of the scatterer velocity.

When t_0 is large enough so that the scatterer velocity varies during the acquisition, an analysis in terms of modulation of the time of flight of the scattered wave shows that the wave is phase modulated. To show this, we will consider the unidimensional situation where a plane progressive monochromatic sound wave propagating at velocity c in a quiescent medium is backscattered by a plane scatterer. This object has an infinite acoustic impedance (normal total reflection) and oscillates sinusoidally around its mean position according to the trajectory $x_S(t)=A\sin\Omega t$. We assume that the wave is emitted at time $t-\tau(t)$ by a transducer located at a distance L from the scatterer. The backscattered wave is then detected by the same transducer at time t such that

$$\tau(t)=\frac{2}{c}\left[L-x_S\left(t-\frac{\tau(t)}{2}\right)\right]. \tag{1}$$

This modulation of the time of flight of the wave induces a phase modulation of the backscattered wave detected at a distance L from the scatterer, such that

$$p^D \propto \exp i\omega(t-\tau(t)), \tag{2}$$

where the superscript "D" denotes the Doppler contribution to the backscattered wave. If $A \ll L$, we have $\tau(t)\approx 2L/c$. In addition, if the velocity of the plate is small compared to the sound velocity c, $M \equiv A\Omega/c \ll 1$, substitution of $\tau(t)\approx 2L/c$ in the argument of the vibration amplitude x_S in (1) is justified and leads to an explicit approximate expression for $\tau(t)$. This gives a phase modulation of the form

$$\omega\tau(t)\approx\frac{2\omega}{c}\left[L-x_S\left(\frac{t-L}{c}\right)\right]. \tag{3}$$

This is what is called the quasi-static approximation of the Doppler effect. The detected wave then rewrites

$$p^D \propto \exp i\left[\omega t+2kA\sin\Omega\left(t-\frac{L}{c}\right)\right], \tag{4}$$

where $k=\omega/c$ is the wave number of the high frequency incident acoustic wave. The generation of sideband peaks at pulsation $\omega\pm n\Omega$ (n integer) in the spectrum of the backscattered wave is evidenced when transforming the latter expression to

$$p^D \propto \exp i(\omega t)\sum_{n=-\infty}^{+\infty} J_n(2kA)\exp\left[in\Omega\left(t-\frac{L}{c}\right)\right], \tag{5}$$

where J_n is the Bessel function of nth order. The sideband peaks at frequency $f\pm nF$ ($F=\Omega/2\pi$) have their amplitude proportional to $|J_n(2kA)|$. For $2kA \ll 1$, $J_n(2kA) \sim (kA)^n/n!$, and the Doppler effect is considered as weak, i.e., the energy of the peak at frequency f in the spectrum of the backscattered wave is almost equal to the energy of the incident wave, and the leading sideband peaks are at frequency $f\pm F$. Note that the condition $M \ll 1$ implies two possible situations: the first is $F \ll f$, thus Eq. (5) is valid for any value of $2kA$ such that $2kA \ll f/F$; the second is $F \sim f$, thus $A\omega/c \sim A\Omega/c \ll 1$, and therefore Eq. (5) is restricted to $kA \ll 1$.

Finally, the general case of oblique incidence has been widely studied theoretically.[1,3,7,19,20] If we define θ as the angle between the incident wave and the normal of the surface, the argument of the Bessel functions of the scattered wave [Eq. (5)] must be replaced by

$$2kA \rightarrow (k_z+k_{nz})A, \tag{6}$$

where $k_z=k\cos\theta$, $k=\omega/c$,

$$k_{nz}=\sqrt{\left(\frac{\omega+n\Omega}{c}\right)^2-k_x^2}, \tag{7}$$

and $k_x=k\sin\theta$. The argument $(k_z+k_{nz})A$ depends then on n and θ. If $F \ll f$ and if we restrict to the first sidebands, we can approximate $(\omega+n\Omega)/c \approx \omega/c$ in (7). Thus, the argument of the Bessel functions results,

$$(k_z+k_{nz})A \approx 2kA\cos\theta, \tag{8}$$

and, for small angles, the multiplicative correction term is almost unity, $\cos\theta \approx 1-\theta^2/2$. Thus, from an experimental point of view, a simple configuration to study is $\theta \ll 1$ and $F \ll f$; in this case, Eq. (5) holds for any value of $2kA$ such that $2kA \ll f/F$.

B. Surface quasi-static Doppler effect versus bulk nonlinear effects

Due to the intrinsic nonlinear character of the conservation equations and to the nonlinear dependence of pressure fluctuations against density fluctuations (as a consequence of the equation of state of the fluid), sound propagation is nonlinear.[4,5] Thus, two collinear waves of frequencies f_1 and f_2 may interact and generate waves whose frequencies are linear combinations of f_1 and f_2, and whose amplitudes increase with the distance of interaction L.[5] Considering our experiment, the nonlinear interaction of the low frequency wave p_Ω emitted by the vibrating scatterer and the high frequency backscattered wave p_ω leads to the generation of

sideband peaks in the spectrum of the detected wave. In the case of weak nonlinear interaction, the amplitude of the first sideband is to leading order proportional to

$$p^{NL}_{\omega\pm\Omega}=\frac{\epsilon}{2\rho c^3}(\omega\pm\Omega)p_\Omega p_\omega L, \tag{9}$$

where $\epsilon=B/2A+1$ is the usual nonlinear parameter of the medium[5,21] (here A should not be confused with the vibration amplitude). For gases, $\epsilon=(\gamma+1)/2$, with $\gamma=c_p/c_v$ the specific heat ratio.

Therefore, when an acoustic wave is scattered by a vibrating surface, both the surface Doppler effect and bulk non-linearities contribute to the generation of combination frequencies. However, it has been argued that the Doppler effect is practically always dominated by bulk non-linearities.[6,8–11] In the case of plane waves, Piquette and Van Buren[10] and later Bou Matar *et al.*[18] predicted that when the dimensionless parameter

$$Y=\frac{p^D_{\omega\pm\Omega}}{p^{NL}_{\omega\pm\Omega}}=\frac{2\rho c^2 A}{\epsilon L p_\Omega} \tag{10}$$

is large (resp. small) compared to unity, the Doppler effect (resp. bulk nonlinear effect) is dominant. Note that in addition to the plane wave assumption, such that $p_\Omega=\rho c v_\Omega$, these studies use the approximation that the amplitude of the low frequency velocity field is given by the scatterer velocity, i.e., $v_\Omega\approx A\Omega$, obtaining $Y=\Lambda/\epsilon\pi L$, where $\Lambda=c/F$. Completing the experimental demonstration by Bou Matar[18] of the existence of two asymptotic regimes, a Doppler dominant regime and a bulk nonlinearity dominant one, we showed in our previous experimental study[12] that the criterium $Y=1$, expressed in its more general form Eq. (10), is quantitatively correct.[22] In this article, we focus on the Doppler effect, and therefore we choose an experimental configuration where the condition $Y\gg 1$, as defined by Eq. (10), is verified.

III. EXPERIMENTAL SETUP

Figure 1 displays the experimental setup. The vibrating scatterer consists of a square flat piston made of PMMA, 17×17 cm and 10 mm thick, located at $x=0$ say. It is driven sinusoidally by an electromechanical vibration exciter of Bruel & Kjaer (BK) 4808 type, at a frequency $F=14$ Hz and amplitude $10^{-6}\,\mathrm{m}<A<3.5\times10^{-3}\,\mathrm{m}$. An air coupled transducer ITC 9073, $d=12$ mm in diameter, located at a distance $L=28$ cm, generates a wave at frequency $f=225$ kHz, incident on the vibrating plane with a small angle $\theta=5°$. It is driven by the source of a high frequency spectrum analyzer Agilent 3589 amplified by a NF Electronic Instruments 4005 power amplifier. The incident wave on the piston is in the far field, as $d^2/4\lambda\approx 23.5\,\mathrm{mm}\ll L$. The backscattered wave is detected by another ITC 9073 transducer, also located at a distance L, and oriented with an angle $\theta=-5°$. The displacement of the piston is controlled by a Wavetek 395 function generator. This sinusoidal electric signal is amplified by a BK 2712 power amplifier, and the displacement is of the form $x_S(t)=A\sin(\Omega t)$. We note that the vibration harmonics have an energy at least 100 times smaller that the fundamental frequency vibration. The acceleration $-A\Omega^2\sin(\Omega t)$ of the scatterer is measured using a BK 4393V piezoelectric accelerometer and a BK 2635 charge amplifier. The acceleration signal is processed using both an Agilent 35670A low frequency spectrum analyzer and a Stanford Research System 830 Lock-In amplifier, in order to get the displacement of the scatterer at F (the difference between both measurements is smaller than 0.5%). The power spectrum of the backscattered wave is computed with the high frequency Agilent 3589A spectrum analyzer. The experiment is controlled by a Power PC Mac computer and the data are transferred to this computer via a general purpose interface bus (GPIB) board.

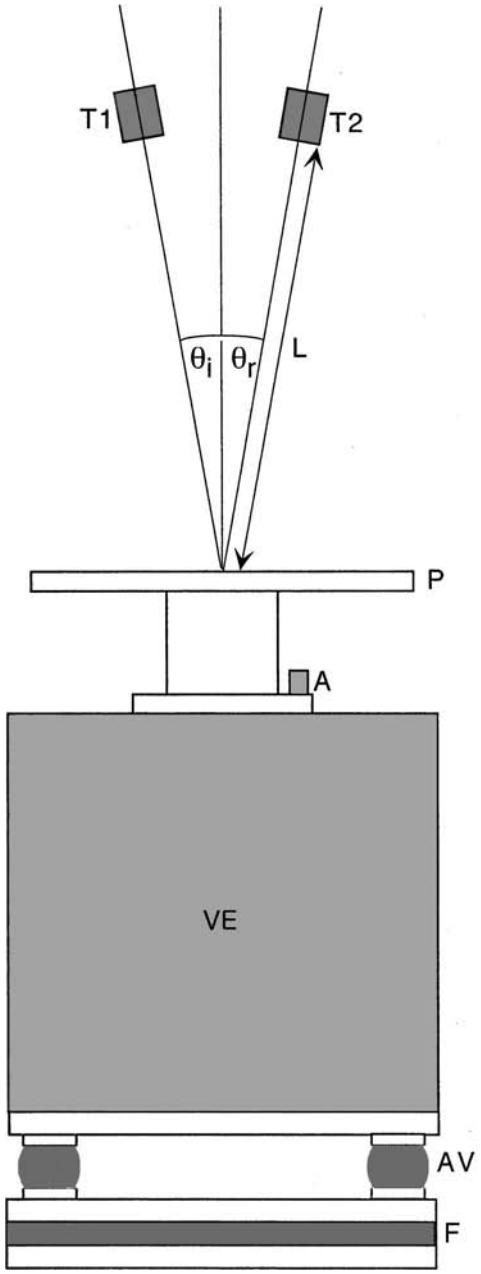


FIG. 1. Sketch of the experimental apparatus. (VE) electromechanical vibration exciter, (P) 17×17×1 cm PMMA piston, (A) piezoelectric accelerometer, (T1-T2) air coupled transducers, (AV) anti-vibrations supports, and (F) foam anti-vibration layer. In the experiments described here, $\theta_i=\theta_r\equiv\theta=5°$ and $L=280$ mm. The piston displacement is $x_S(t)=A\sin(\Omega t)$.

Concerning the power spectrum measurements, an important parameter is t_0F, where t_0 is the time duration of the analyzed time series, and t_0^{-1} is thus the spectral resolution (SR). In practice, it is the frequency span, f_{sp}, of the Agilent spectrum analyzer which is varied for a fixed number of points, $N=401$. Thus, $t_0^{-1}=f_{sp}/(N-1)$, and the frequency span takes the values 5 kHz, 2.5 kHz, 1.25 kHz, 625 Hz, and 312.5 Hz, giving $t_0^{-1}=12.5$, 6.25, 3.125, 1.5625, and 0.78125 Hz.

Finally, in our experiment $\theta=5°$ is finite but small. As discussed in Sec. II, if $F \ll f$ and if we restrict our analysis to the first sidebands, the argument of the Bessel functions appearing in Eq. (5) must be replaced by $2kA\cos\theta$ [see Eq. (8)]. This is indeed the case in our experiments, since $F \sim O(10\,\text{Hz})$, $f \sim O(225\,\text{kHz})$, and the number of analyzed sidebands is either 5 or 11. For $\theta=5°$, the correct argument of the Bessel functions is $2kA\cos\theta \approx 1.992kA$, which represents a 0.4% error with respect to $2kA$. This error is small compared to the errors related to the measurements of A and the error on k (given by the approximation $c \approx 345 \pm 5$ m/s at 25 °C), and it is then considered as negligible. Therefore, we simply use the parameter $2kA$ in the analysis of our experimental results.

IV. MAIN CHARACTERISTIC FEATURES OF THE DOPPLER EFFECT

A. Static Doppler effect

As explained above, when $t_0^{-1} \gg F$ the time averaged backscattered wave power spectrum is expected to be proportional to the PDF of the scatterer velocity. In addition, a static Doppler shift can be detected only if it exceeds the spectral resolution. This implies an additional condition, $2fA\Omega/c > t_0^{-1}$, i.e., $2kA \gg 1$. In our experimental setup both conditions are simultaneously verified for $F=5$ Hz and $A \geqslant 1.7$ mm. In addition $Y \gg 1$, and we are then in the Doppler dominated regime.

Figure 2(a) displays the backscattered wave power spectrum density (PSD) obtained for a low spectral resolution (SR) of $t_0^{-1}=12.5$ Hz, such that $t_0^{-1} \approx 2.5F$, and a scatterer vibration amplitude A verifying $2kA=32.6$. Thus, the conditions for observing a static Doppler shift are approximately verified. We observe that the spectrum displays two maxima, and we define Δf as the frequency shift of these maxima; in this case $\Delta f \approx 150$ Hz. These maxima can be identified with the two maxima of the scatterer velocity PDF occurring at $\pm A\Omega$. Its theoretical expression

$$\text{PDF}[\dot{x}_S/A\Omega] = \frac{1}{\pi \sin(\arccos(\dot{x}_S/A\Omega))} \tag{11}$$

is presented in Fig. 2(b) together with the experimental PDF of the actual scatterer velocity normalized by the value $A\Omega = 0.125$ m/s, obtained with the acceleration measurement done with the lock-in amplifier. Note that in Fig. 2(a) two pairs of extra symmetric peaks appear at roughly $2\Delta f$ and $3\Delta f$, but they are of much lower intensity. They probably correspond to the Doppler shifts of the waves which undergo multiple reflections on the scatterer.

Figure 3 displays the variations of Δf as a function of the expected static Doppler frequency shift $2fA\Omega/c$. The solid line represents the expected correlation between these two quantities in the static Doppler effect regime. The systematic underestimation of Δf is probably due to the fact that the scatterer velocity is not constant during the time of signal acquisition. Nevertheless, there is a satisfactory correlation between both quantities.

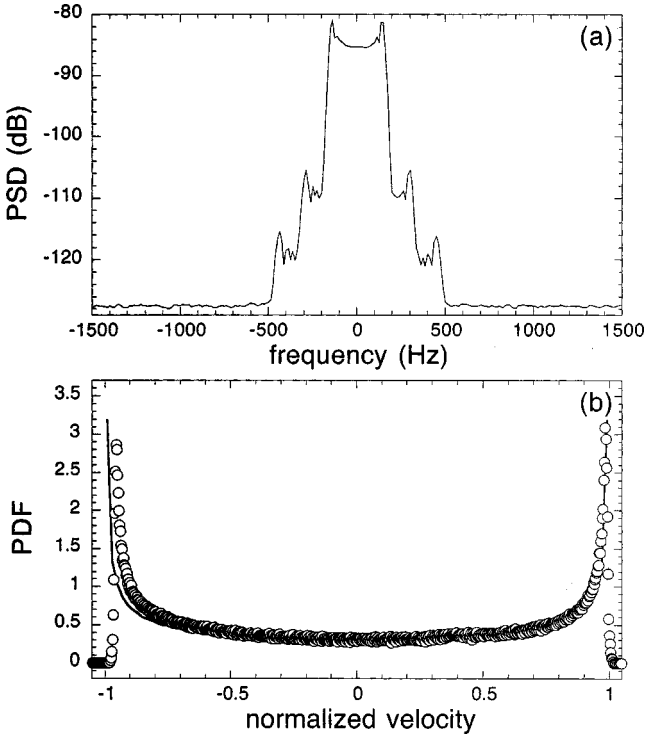


FIG. 2. (a) Backscattered wave power spectrum density (PSD) for $F=5$ Hz and $t_0^{-1}=12.5$ Hz. Frequencies are shifted by $f=225$ kHz. (b) Experimental (○) and theoretical (solid line) PDFs of the normalized scatterer velocity.

Figure 4 displays the backscattered wave power spectra

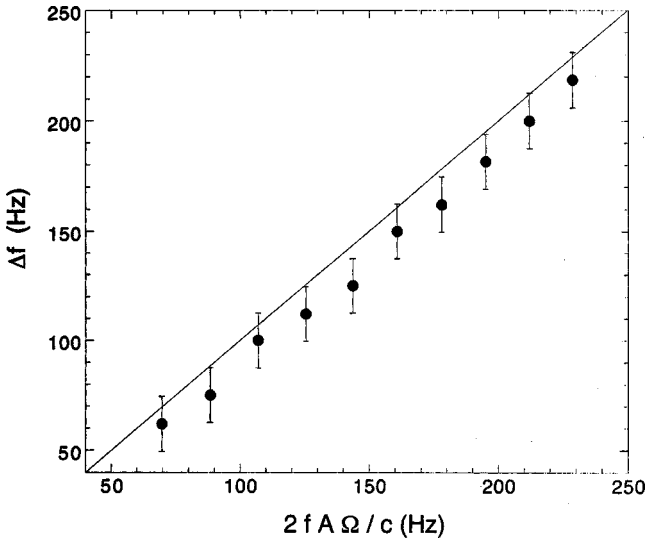


FIG. 3. Maximal frequency shift measured from the backscattered wave power spectrum Δf as a function of $2A\Omega f/c$ which is the expected static Doppler shift obtained from acceleration measurements. Error bars correspond to the frequency resolution $t_0^{-1}=12.5$ Hz and $F=5$ Hz. Solid line represents $\Delta f = 2A\Omega f$.

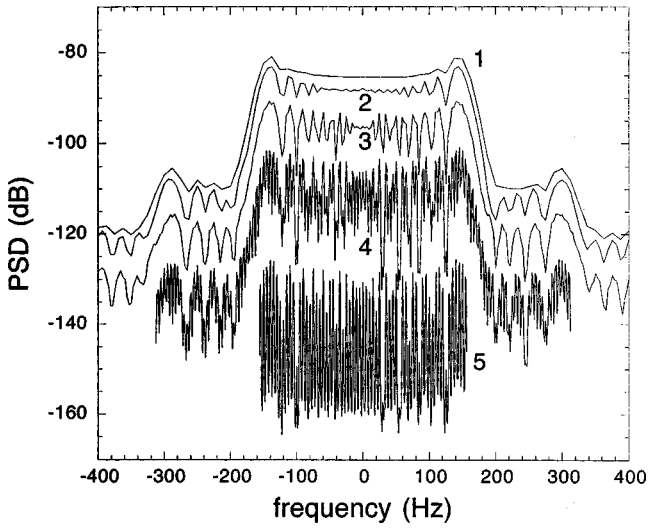


FIG. 4. Backscattered wave power spectra for $F=5$ Hz and $t_0^{-1}=12.5$ Hz (1), 6.25 Hz (2), 3.125 Hz (3), 1.5625 Hz (4), and 0.78125 Hz (5). Frequencies were shifted by $f=225$ kHz. Curves (3)–(5) have been shifted down by 5, 15, and 40 dB, respectively, for a better display.

obtained for various spectral resolutions ranging from $t_0^{-1}=12.5$ to 0.78125 Hz. For a larger SR, $t_0^{-1}=6.25\,\mathrm{Hz}\gtrsim F$, the condition of static Doppler effect is not verified, and curve 2 of Fig. 4 displays several peaks between the two maxima of curve 1. The number of peaks increases with the spectral resolution (curves 3 and 4) until the frequency difference between two peaks drops to $F=5$ Hz (curve 5). As we explain in the following, this can be understood in the frame of the quasi-static Doppler effect presented in Sec. II. Note that whereas in the static Doppler effect regime the frequency difference between two peaks in the scattered wave power spectrum is determined by the scatterer oscillation *velocity amplitude*, in the quasi-static Doppler effect regime it is determined by the scatterer oscillation *frequency*.

B. Quasistatic Doppler effect

In this section we present experimental results concerning the quasi-static Doppler effect. We show that in our current experimental configuration, which is slightly different than in Ref. 12, the Doppler effect overcomes bulk nonlinearities in the generation of the sidebands at $f\pm nF$ of the backscattered wave. Although the prediction of the phase modulation theory concerning the amplitude of the first sideband is often used in the "weak regime" $2kA\ll 1$ in various applications,[13–18] the phase modulation theory has never been experimentally verified, especially in the "strong regime" $2kA\gtrsim 1$.

Three pressure power spectra of the backscattered wave computed for $f=225\,\mathrm{kHz}$, $F=14\,\mathrm{Hz}$, and three different values of the dimensionless amplitude $2kA$ are displayed in Fig. 5. (For the calculation of $k=\omega/c$ we used $c\approx 345$ m/s for air at ambient temperature $T_0=25\pm 1\,^\circ$C.) These spectra are composed of pairs of symmetric peaks at $f\pm nF$ surrounding the central peak at *f*. Comparison of spectra 5a and b shows that the number of sidebands emerging from the background noise around the central peak at *f* increases with $2kA$, exceeding the frequency span for $2kA\gtrsim 1$. A first cancellation of the central peak occurs at $2kA\approx 2.3$, as shown in Fig. 5(c), where the central peak has decreased by approximately 30 dB with respect to its initial value.

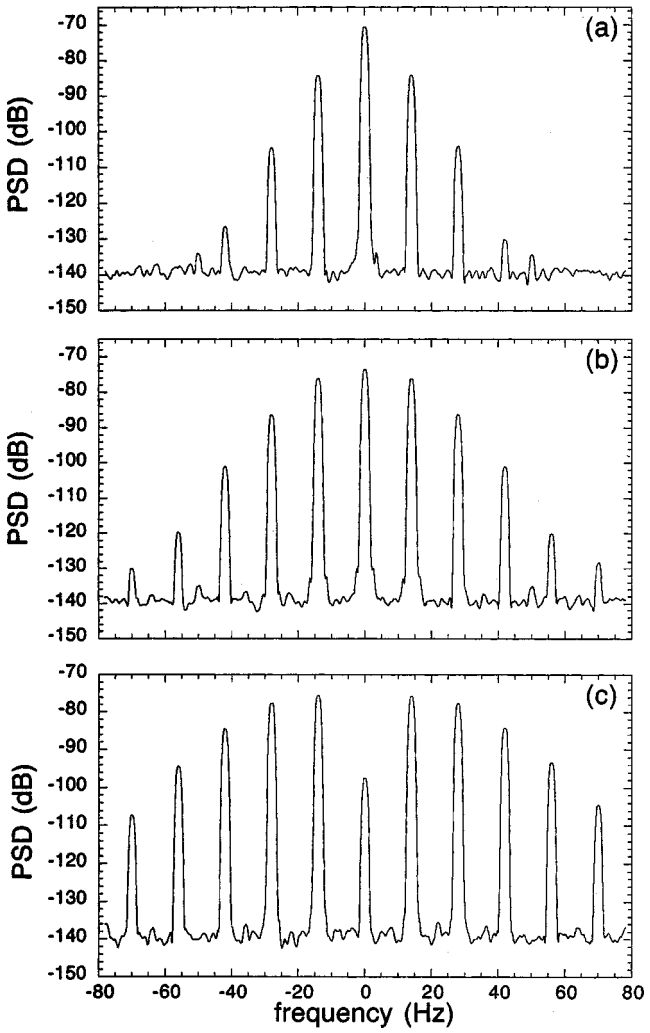


FIG. 5. Power spectra of the wave scattered by the vibrating plate as a function of the dimensionless vibration amplitude $2kA=0.39$ (a), 1.17 (b), and 2.34 (c) (k is the wave number of the incident wave of high frequency *f*). Frequencies were shifted by $f=225$ kHz. For the calculation of $k=\omega/c$ we used $c=345$ m/s for air at ambient temperature $T_0=25\pm 1\,^\circ$C.

Figure 6 displays the variations of the normalized amplitudes $p_n=p_{\omega+n\Omega}/p_{\omega 0}$ of the side-bands measured in the power spectrum of the backscattered wave as a function of $2kA$, in both the weak (a), $2kA\ll 1$, and strong (b,c), $2kA\gtrsim 1$, regimes. Normalization is done using the amplitude $p_{\omega 0}$ of the wave backscattered on the motionless piston (called hereafter the reference wave). Comparison with the predictions of phase modulation theory, i.e., with the absolute values of the Bessel functions, shows excellent agreement, except for high order sidebands when their amplitude is close to the background noise (this is for $p_n\lesssim -65\,\mathrm{dB}$, i.e., $p_n\lesssim 6\times 10^{-4}$). For $2kA$ small, the amplitude of the sidebands at frequency $f\pm nF$ $(n=0-5)$ increases like $(kA)^n/n!$. For higher values of $2kA$, the amplitude of the component at frequency *f* of the scattered wave decreases and we observe that it almost vanishes for $2kA\approx 2.3$ [see Figs. 5(c) and 6(b)]. It then increases when $2kA$ is increased further above 2.3, and then decreases and vanishes again for $2kA\approx 5.4$.

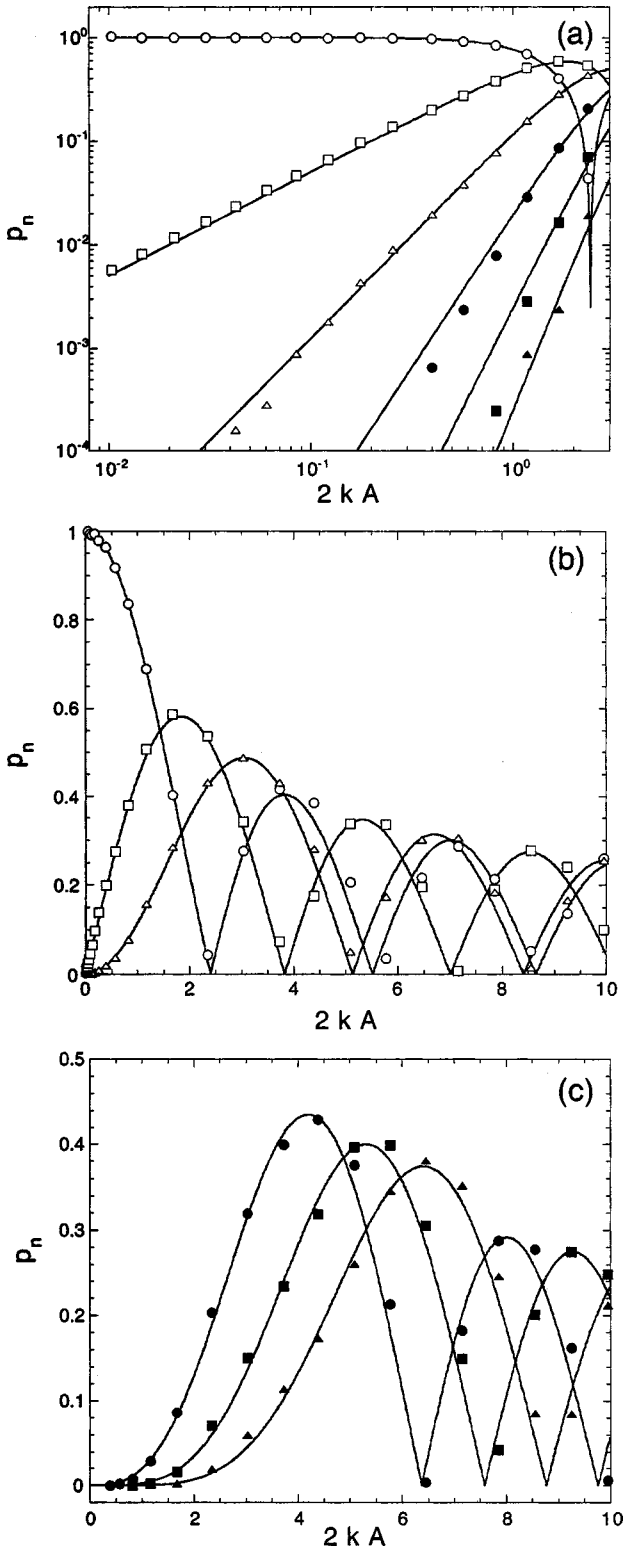


FIG. 6. Normalized pressure $p_n = p_{\omega+n\Omega}/p_{\omega 0}$ versus $2kA$ for $n=0$ (○), 1 (□), 2 (△), 3 (●), 4 (■), and 5 (▲). $p_{\omega 0}$ is the amplitude of the wave backscattered on the motionless piston. (a) Weak regime in $\log_{10}-\log_{10}$ scale and (b) and (c) strong regime in linear scale. Continuous lines show the absolute values of the Bessel functions of order $n=0,...,5$. Note that there is no adjustable parameter, and k is fixed by the sound speed c.

The same process occurs roughly periodically for the amplitude of each peak $f \pm nF$, as displayed in Fig. 6 for $n=0-5$. We note that experiments were performed up to $2kA=30$, showing the same behavior.

Another way to check that the Doppler effect is dominant versus bulk acoustic nonlinearities is to vary the vibration frequency, always keeping $Y \gg 1$, and to show that the sideband amplitudes p_n scale like A, independently of F. This is indeed the case in our experimental configuration. Thus, the Doppler effect is the dominant mechanism in the generation of frequency combinations $f \pm nF$. We can then study an acoustic nonintrusive estimator of the vibration amplitude A, which is presented in the next section.

V. TEST OF A SPECTRAL ESTIMATOR OF THE SCATTERER HARMONIC VIBRATION AMPLITUDE

Recently, Huang *et al.*[15] proposed to use a mathematical property of Bessel functions,

$$z = \sqrt{\frac{2\Sigma_{n=-\infty}^{+\infty} n^2 J_n^2(z)}{\Sigma_{n=-\infty}^{+\infty} J_n^2(z)}}, \tag{12}$$

in order to define a spectral estimator $\hat{A}$ of the scatterer vibration amplitude A when its motion is harmonic in the quasi-static Doppler regime. In practice, the number of detected sidebands being finite, the estimator has to be defined with a finite number of sidebands $2n_0$:

$$\hat{A} = \frac{1}{\sqrt{2}k} \sqrt{\frac{\Sigma_{n=-n_0}^{+n_0} n^2 p_n^2(2kA)}{\Sigma_{n=-n_0}^{+n_0} p_n^2(2kA)}}. \tag{13}$$

We can then test the accuracy and validity domain of this estimator. It is expected to be valid as long as the energy of the undetected sidebands is small compared to the energy of the detected ones. This can be checked by measuring the energy of the backscattered wave E and comparing it to the energy of the reference wave E_0. As a matter of fact, due to the orthogonality of the Bessel functions, the normalized energy of a wave of the form given by Eq. (5) is

$$\frac{E}{E_0} = \sum_{n=-\infty}^{+\infty} J_n^2(2kA) = 1. \tag{14}$$

As long as the undetected energy is small compared to the one of the reference wave, we expect the estimator $\hat{E}$ to be equal to the energy of the reference wave:

$$\hat{E} = \sum_{n=-n_0}^{+n_0} p_n^2(2kA) \approx E_0. \tag{15}$$

This reflects the fact that the phase modulation process induced by the scatterer vibration does not transfer any energy from the scatterer to the scattered wave. In fact, the energy of the sidebands is pumped from the incident wave only.

Figure 7(a) displays the average relative error $\hat{A}/A-1$ as a function of $2kA$ for $n_0=5$ and 11, where A is determined from the scatterer acceleration. The averages are performed over 22 independent experimental runs for each value of A. This average relative error is compared to the error computed using a theoretical truncated amplitude estimator $\hat{A}_T$, which is defined with the finite number of detected sidebands:

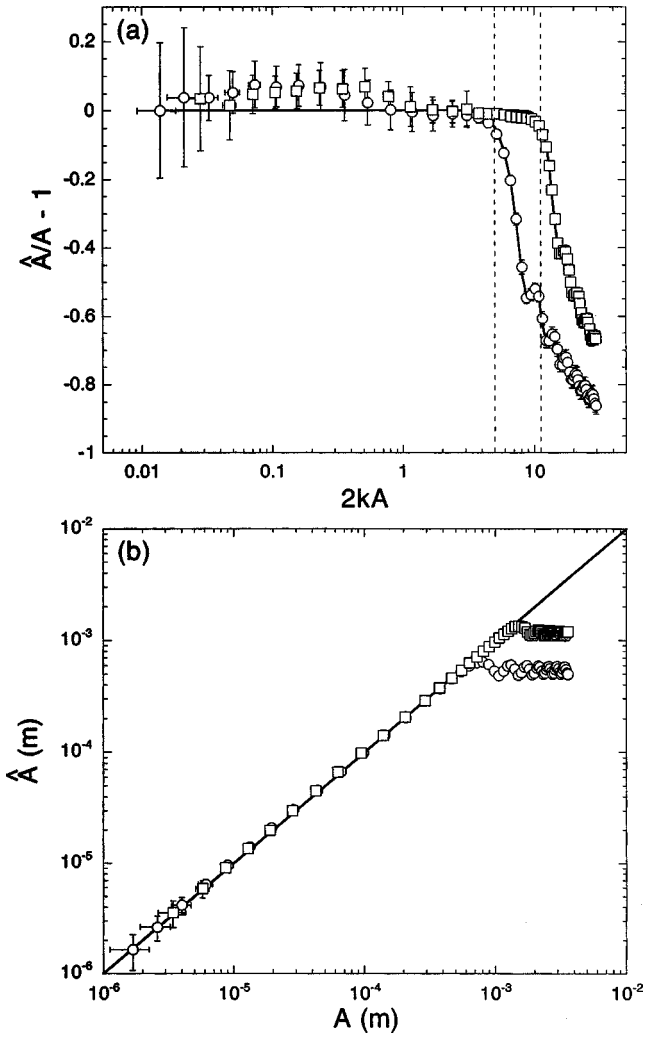


FIG. 7. (a) Amplitude estimator error, $\hat{A}/A-1$, versus $2kA$. The solid line represents the theoretical truncated amplitude estimator error. Vertical dashed lines represent the values $2kA=n_0$. Error bars correspond to the errors obtained including the standard deviation of 22 independent experimental runs, the errors associated to the amplitude measurements, and the error associated to the value of k ($\pm 1.5\%$ for c). (b) Estimated amplitude $\hat{A}$ versus A. In this case the error bars correspond to the statistical standard deviation of 22 independent experimental runs. In both figures, results for two PSD spans are presented, $n_0=5$ (○) and 11 (□).

$$\hat{A}_T=\frac{1}{\sqrt{2}k}\sqrt{\frac{\Sigma_{n=-n_0}^{+n_0} n^2 J_n^2(2kA)}{\Sigma_{n=-n_0}^{+n_0} J_n^2(2kA)}}. \tag{16}$$

In the range $2kA<n_0$, the estimator $\hat{A}$, defined in Eq. (13), is found to have a satisfactory accuracy. The mean value of $\hat{A}/A-1$ does not exceeds 5%. The associated errorbars are important (of the order of 20%) for $2kA\approx 0.01$, but they decrease rapidly to a value below 5% as we increase A. The larger errors are in fact due to the uncertainty of the determination of the amplitude from acceleration measurements at low frequency (At both low F and low A, the acceleration is quite small, of the order of $1.6\times 10^{-2}\,\mathrm{m/s^2}$.) However, the relative errors of both amplitude measurements, i.e., the acceleration and nonintrusive acoustic power spectrum estimator measurements, are of same order of magnitude, as can be seen in Fig. 7(b), where we present $\hat{A}$ versus A in a $\log_{10}-\log_{10}$ scalc.

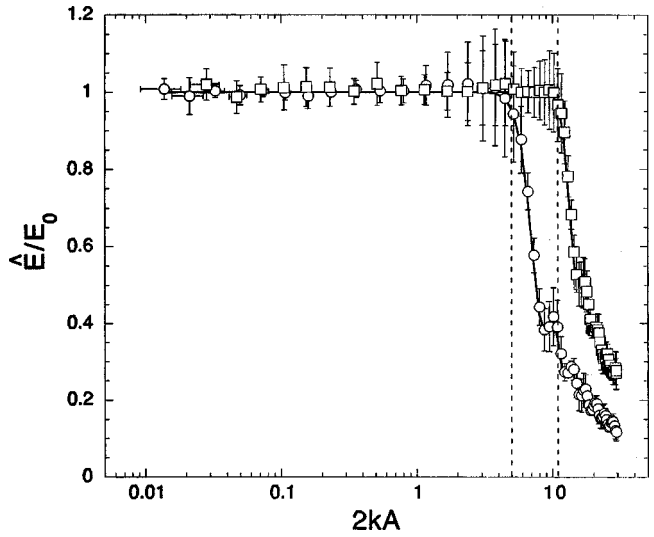


FIG. 8. Estimated normalized acoustic energy $\hat{E}/E_0$ versus $2kA$ for two PSD spans, $n_0=5$ (○) and 11 (□). In this case the errorbars correspond only to the standard deviation of 22 independent experimental runs. Vertical dashed lines represent the values $2kA=n_0$.

In Fig. 7(a) we observe that for $n_0=5$ (resp. 11) the estimation error increases above 5% when $2kA>5$ (resp. 11), in agreement with the error computed using the theoretical truncated amplitude estimator $\hat{A}_T/A-1$. This departure from 0 is concomitant with the saturation of $\hat{A}$ observed in Fig. 7(b). These facts are due to the transfer of a significant part of the incident energy to undetected sidebands during scattering. This is clearly observed in Fig. 8, which shows the average value of $\hat{E}/E_0$ as a function of $2kA$. It is compared to the normalized energy computed using a theoretical truncated energy estimator $\hat{E}_T/E_0$, defined as

$$\frac{\hat{E}_T}{E_0}=\sum_{n=-n_0}^{+n_0} J_n^2(2kA). \tag{17}$$

Here, we also find that both $\hat{E}/E_0$ and $\hat{E}_T/E_0$ drop by more than 5% when $2kA>n_0$.

We thus conclude that for the experimental configuration presented here, the acoustic amplitude estimator proposed in Ref. 15 is valid in a large range of amplitudes, namely $2\times 10^{-2}<2kA<n_0$, which gives $2\times 10^{-3}\,\mathrm{mm}<A<n_0/2k\approx 1.2\,\mathrm{mm}$ (resp. 2.7 mm) for $n_0=5$ (resp. $n_0=11$). The relative error on $\hat{A}$ is of the same order of magnitude as that on the acceleration measurements. The lower limit of A is given by the signal-to-noise ratio (SNR) of the acceleration measurements whereas the lower limit of $\hat{A}$ is given by the PSD SNR, which is of the order of 65 dB in our present experimental configuration. Concerning the acoustic nonintrusive estimator $\hat{A}$, for small amplitude measurements ($2kA\ll 1$) the first sidebands must then satisfy $p_{\omega\pm\Omega}/p_{\omega 0}=p_1/p_0\approx kA>-65\,\mathrm{dB}\approx 6\times 10^{-4}$ in order to be resolved. Thus, to get a better resolution on the measured amplitude we have to increase k, i.e., to increase f.[23] On the other hand, to increase the upper limit of the measurable amplitude, one has to in-

crease the ratio n_0/k, either by increasing the number of sidebands taken into account in the analysis or by decreasing k (by decreasing f).

VI. CONCLUSION

We have studied the Doppler effect generated by vibrating scatterers. We have shown that depending on the duration t_0 of the analyzed backscattered wave, we have two possible situations: (i) if $t_0 \ll F^{-1}$, the static Doppler effect is observed and the backscattered power spectrum is proportional to the PDF of the scatterer velocity. The frequency shift Δf is then given by the amplitude of the scatterer velocity. (ii) If $t_0 \gg F^{-1}$, the so-called quasi-static Doppler effect is observed, and the scattered wave is phase modulated by the normalized vibration displacement $2kx_s(t) = 2kA \sin(\Omega t)$. In this case, the backscattered wave has a frequency spectrum composed of peaks at $f \pm nF$ (n integer). Experiments show that the backscattered wave is very well described by Eq. (5), i.e., the amplitudes of the sidebands of order n are given by the Bessel functions of order n. We verified that in the case $F \ll f$, this agreement holds in both the weak ($2kA \ll 1$) and strong ($2kA \gtrsim 1$) regimes.

The last part of this work was devoted to the study of a nonintrusive acoustic estimator of the amplitude vibration A. We conclude that for the present experimental configuration, the acoustic amplitude estimator is valid for a large range of amplitudes, namely $2 \times 10^{-2} < 2kA < n_0$, which gives $2 \times 10^{-3}\,\text{mm} < A < n_0/2k \approx O(1\,\text{mm})$. To increase the upper bound one has to take into account a larger number of sidebands or one can chose to decrease the incident sound frequency. For a given SNR, one has to increase the incident frequency to decrease the lower bound.

ACKNOWLEDGMENTS

RW thanks the Center National d'Etudes Spatiales (CNES) for financial support. This work has been supported by CNES Contract No. 03/11/21/00.

[1] A. M. Kovalev and V. N. Krasil'nikov, "Reflection of electromagnetic waves from moving surfaces," Sov. Phys. Tech. Phys. **7**, 19 (1962).
[2] D. Censor, "Scattering by time varying obstacles," J. Sound Vib. **25**, 101–110 (1972).
[3] D. Censor, "Acoustical Doppler effect analysis—Is it a valid method?" J. Acoust. Soc. Am. **83**, 1223–1230 (1988).
[4] M. F. Hamilton and C. L. Morfey, "Model equations," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, New York, 1998).
[5] K. Naugolnykh and L. Ostrovsky, *Nonlinear Wave Processes in Acoustics* (Cambridge U.P., Cambridge, 1998).
[6] P. H. Rogers, "Comments on 'Scattering by time varying obstacles'," J. Sound Vib. **28**, 764–768 (1973).
[7] D. Censor, "Harmonic and transient scattering from time varying obstacles," J. Acoust. Soc. Am. **76**, 1527–1534 (1984).
[8] J. C. Piquette and A. L. Van Buren, "Nonlinear scattering of acoustic waves by vibrating surfaces," J. Acoust. Soc. Am. **76**, 880–889 (1984).
[9] J. C. Piquette and A. L. Van Buren, "Comments on Harmonic and transient scattering from time varying obstacles [J. Acoust. Soc. Am. **76**, 1527–1534 (1984)]," **79**, 179–180 (1986) and author's reply.
[10] J. C. Piquette and A. L. Van Buren, "Some further remarks regarding scattering of an acoustic wave by a vibrating surface," J. Acoust. Soc. Am. **80**, 1533–1536 (1986).
[11] J. C. Piquette, A. L. Van Buren, and P. H. Rogers, "Censor's acoustical Doppler effect analysis—Is it a valid method?" J. Acoust. Soc. Am. **83**, 1681–1682 (1988).
[12] N. Mujica, R. Wunenburger, and S. Fauve, "Scattering of a sound wave by a vibrating surface," Eur. Phys. J. B **33**, 209–213 (2003).
[13] M. Cox and P. H. Rogers, "Automated noninvasive motion measurement of auditory organs in fish using ultrasound," J. Vib. Sound Str. Rel. Design **109**, 55–59 (1987).
[14] R. M. Lerner, K. J. Parker, J. Holen, R. Gramiak, and R. C. Waag, "Sono-elasticity: medical elasticity images derived from ultrasound signals in mechanically vibrated targets," Acoust. Imaging **16**, 31–37 (1988).
[15] S. Huang, R. M. Lerner, and K. J. Parker, "On estimating the amplitude of harmonic vibration from the Doppler spectrum of reflected signals," J. Acoust. Soc. Am. **88**, 2702–2712 (1990).
[16] Y. Yamakoshi, J. Sato, and T. Sato, "Ultrasonic imaging of internal vibration of soft tissue under forced vibration," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**, 45–53 (1990).
[17] F. Figueroa and E. Barbieri, "An ultrasonic ranging system for structural vibration measurements," IEEE Trans. Instrum. Meas. **40**, 764–769 (1991).
[18] O. Bou Matar, J. P. Remenieras, C. Brunee, A. Roncin, and F. Patat, "Noncontact measurement of vibration using airborne ultrasound," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**, 626–633 (1998).
[19] L. M. Lyamshev, "Diffusion of sound by a periodically moving plate," Sov. Phys. Dokl. **28**, 274–276 (1983).
[20] L. M. Lyamshev, "Distinctive features of the scattering and radiation of sound by periodically moving plates and shells," Sov. Phys. Acoust. **30**, 237–238 (1984).
[21] R. T. Beyer, "The parameter B/A," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, New York, 1998).
[22] The Doppler effect can dominate nonlinear effects for $Y < 1$ in the case of diverging waves because of a less efficient nonlinear interaction in the bulk whereas the Doppler effect remains unchanged.[12] The opposite effect is expected for converging waves.
[23] For example, for experiments in water presented in Ref. 12, where the SNR is 80 dB, $L = 20$ cm, $F = 120$ Hz, and $f = 2.25$ MHz, using only one sideband, $n_0 = 1$, and for a single experimental run, we find that the relative error $\hat{A}/A - 1$ is lower than 15% for $7 \times 10^{-9}\,\text{m} < A < 5 \times 10^{-5}\,\text{m}$. Smaller errors can be obtained by averaging over several experimental runs as is done the present work.

Scattering of antiplane shear waves by layered circular elastic cylinder

Liang-Wu Cai
Department of Mechanical and Nuclear Engineering, 331 Rathbone Hall, Kansas State University, Manhattan, Kansas 66506

(Received 26 March 2003; revised 1 November 2003; accepted 6 November 2003)

An exact analytical solution for the scattering of antiplane elastic waves by a layered elastic circular cylinder is obtained. The solution and its degenerate cases are compared with other simpler models of circular cylindrical scatterers. The effects of the geometrical and physical properties of the interphase are studied. Numerical results confirm the existence of a resonance mode in which the scatterer's core undergoes a rigid-body motion when the outer layer of the scatterer is very compliant. This resonance mode has been attributed [Liu *et al.*, Science **289**, 1734 (2000)] to a new mechanism for the band gap formed in the extremely low frequency range for phononic crystals made of layered spherical scatterers. Numerical results also show the existence of a similar resonance mode when the outer layer of the scatterer has very high mass density. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1636465]

PACS numbers: 43.20.Gp, 43.20.Fn, 43.40.Fz [GCG] Pages: 515–522

I. INTRODUCTION

Interests in the effects of the interface between the fiber and the matrix in a fiber-reinforced composite material have a long history in the research efforts in composite materials (see reviews by Yim and Williams, 1995a and by Bogen and Hinders, 1994). In these efforts, a fiber is often modeled as a layered circular elastic cylinder, with the outer layer representing the intermediate material phase that is introduced physically or chemically between the fiber and the matrix. Such a model is also suitable for analyzing a fiber that naturally possesses a layered structure, such as a silicon–carbide (SiC) fiber.

As a micromechanics model for composites, scattering by a single fiber often serves as the basis of analysis of problems concerning multiple fibers. A single layered scatterer subjected to a plane-strain wave, often referred to as the *P*/*SV* wave (longitudinal wave, and often referred to as the vertically polarized shear wave in the context of seismic waves), has been studied by several authors, such as Yim and Williams (1995a, 1995b), Sinclair and Addison (1993), and Bogen and Hinders (1993). These analyses result in linear equation systems, which must be solved numerically, for the wave expansion coefficients of various wave fields. Addison and Sinclair (1992) compared the measured ultrasonic response spectra of a single SiC fiber embedded in a Ti matrix with computed results. Chu and Rokhlin (1992) measured several bulk properties of a SiC/Si_3N_4 composite. Huang *et al.* (1995) considered a three-layer SiC fiber-interphase model.

For the case of antiplane shear wave, often referred to as *SH* wave (horizontally polarized wave in the context of seismic waves), only a multilayer model has been considered. Mal and Yang (1994) gave a computational procedure to obtain a solution, while Shindo and Niwa (1996) gave a recursive analytical solution. Mal and Yang (1994) also incorporated a multilayer model into a multiple-scattering analysis to predict bulk properties of composites containing randomly dispersed particles. The author was the first to have studied (Cai, 1998) the exact analytical solution for the dual-layer cylindrical scatterer in his study of multiple scattering of elastic waves in fiber-reinforced composite materials (Cai and Williams, 1999a, 1999b, 1999c).

The applications of such a dual-layer cylindrical model are not limited to modeling composites. At a larger scale, such a model can be modified to analyze cylindrical shell structures found in many civil engineering applications such as underground pipeline and cable systems, in which the *SH* wave case has been of major concern from an earthquake engineering perspective. Furthermore, since the equation that governs elastic *SH* waves and acoustic waves is the same, extending such a model to the field of acoustics is straightforward.

In a recent study by Liu *et al.* (2000) into the acoustic band gap phenomenon, a structure was designed by arranging silicone-coated small lead balls in cubic array in analogy to a crystal structure; a new band gap was identified at a frequency range that is significantly lower than the crystal structure would predict. A new mechanism in the layered spherical scatterer had been attributed to this new phenomenon: acoustic energy was absorbed by the resonance between the scatterer core and its springy coating. Structures could be designed based on this new discovery to permit constructing the acoustic filter in the size of a fraction of the wavelength.

This new discovery, and its promise of novel applications, prompted the author to reexamine the scattering of single layered scatterer in greater detail. The purpose is to study whether the resonance that has been attributed to band gap in the extremely low frequency exists in a two-dimensional model. Furthermore, understanding of the scattering characteristics of individual scatterers is vitally important for the understanding of the scattering process by multiple scatterers, such as in the case of band gap materials.

In this paper, an exact analytical solution for such a

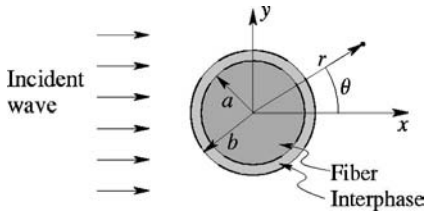


FIG. 1. Problem configuration.

model subjected to an arbitrary incident *SH* wave is derived. The solution is first compared with other simpler models of circular cylindrical scatterers. Then, the effects of the geometrical and physical properties of the interphase are studied through numerical examples.

II. PROBLEM DEFINITION

Consider a layered circular elastic cylinder embedded in a linearly elastic medium of infinite extent, as sketched in Fig. 1. A time-harmonic incident *SH* wave normally impinges upon the scatterer, and the resulting steady-state waves are sought.

Denote the inner and outer radii of the cylinder as a and b $(b>a)$, respectively. Following the terminology in the fiber-interphase-matrix model for composite materials (Yim and Williams, 1995a), the three phases of materials in the model are identified as the fiber, the interphase, and the matrix, respectively. Parameters associated with these phases are signified by the subscripts *f*, *i*, and *m*, respectively. All material properties are assumed to be known.

For an *SH* wave scattering problem, the only nontrivial displacement component is the out-of-plane displacement. Using the polar coordinate system shown in Fig. 1, in the steady state, the nontrivial displacement component is the out-of-plane component, and can be written as

$$w=\phi(r,\theta)\mathrm{e}^{-\hat{i}\omega t}, \tag{1}$$

where $\phi(r,\theta)$ is the complex amplitude satisfying the following Helmholtz equation (see, e.g., Pao and Mow, 1971):

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

where k is the wave number for shear waves, ω is circular frequency, and $\hat{i}=\sqrt{-1}$. Equation (1) signifies that complex notation is used. In complex notation, the physical quantity is either the real or the imaginary part of the product of the quantity's complex amplitude and the temporal factor $\mathrm{e}^{-\hat{i}\omega t}$, and the amplitude of the physical quantity equals the modulus of its complex amplitude.

The complex amplitudes of nonzero stress components can be obtained from that of the displacement as

$$\sigma_{rz}=\sigma_{zr}=2\mu\frac{\partial\phi}{\partial r},\quad \sigma_{\theta z}=\sigma_{z\theta}=\frac{2\mu}{r}\frac{\partial\phi}{\partial\theta}, \tag{3}$$

where μ is the shear modulus.

The general solution to the Helmholtz equation in a polar coordinate system is linear combinations of cylindrical wave functions (Pao and Mow, 1971), which consist of Bessel functions of various kinds as the radial factor and simple harmonics as the angular factor. For the problem at hand, the Helmholtz equation must be satisfied by waves in all three media, and the displacement and the surface traction must be continuous across the interfaces.

Without loss of generality, the incident wave is expressible as (Pao and Mow, 1971)

$$\phi^{\mathrm{inc}}=\sum_{n=-\infty}^{\infty}A_nJ_n(k_mr)\mathrm{e}^{\hat{i}n\theta}. \tag{4}$$

The resulting waves in the three phases are expressible as (Pao and Mow, 1971)

$$\phi^{\mathrm{s}}=\sum_{n=-\infty}^{\infty}B_nH_n^{(1)}(k_mr)\mathrm{e}^{\hat{i}n\theta}, \tag{5}$$

$$\phi^{\mathrm{f}}=\sum_{n=-\infty}^{\infty}C_nJ_n(k_fr)\mathrm{e}^{\hat{i}n\theta}, \tag{6}$$

$$\phi^{\mathrm{i}}=\sum_{n=-\infty}^{\infty}[D_nH_n^{(1)}(k_ir)+E_nH_n^{(2)}(k_ir)]\mathrm{e}^{\hat{i}n\theta}, \tag{7}$$

where ϕ^{s} is the scattered wave in the matrix, ϕ^{f} is the refracted wave in the fiber, ϕ^{i} is the refracted wave in the interphase, $J_n(\cdot)$ is a Bessel function of the first kind, and $H_n^{(1)}(\cdot)$ and $H_n^{(2)}(\cdot)$ are Hankel functions of the first and second kinds, respectively. The total wave in the matrix consists of the incident wave plus the scattered wave; that is,

$$\phi^{\mathrm{total}}=\phi^{\mathrm{inc}}+\phi^{\mathrm{s}}. \tag{8}$$

The boundary conditions are that the stress component σ_{rz} and the displacement ϕ are continuous across both interfaces at $r=a$ and $r=b$. That is

$$(\phi^{\mathrm{inc}}+\phi^{\mathrm{s}})|_{r=b}=\phi^{\mathrm{i}}|_{r=b}, \tag{9}$$

$$2\mu_m\left(\frac{\partial\phi^{\mathrm{inc}}}{\partial r}+\frac{\partial\phi^{\mathrm{s}}}{\partial r}\right)\bigg|_{r=b}=2\mu_i\frac{\partial\phi^{\mathrm{i}}}{\partial r}\bigg|_{r=b}, \tag{10}$$

$$\phi^{\mathrm{i}}|_{r=a}=\phi^{\mathrm{f}}|_{r=a}, \tag{11}$$

$$2\mu_i\frac{\partial\phi^{\mathrm{i}}}{\partial r}\bigg|_{r=a}=2\mu_f\frac{\partial\phi^{\mathrm{f}}}{\partial r}\bigg|_{r=a}, \tag{12}$$

where the expression for σ_{rz} in Eqs. (3) has been used.

Substituting the wave expressions in Eqs. (4) through (7) into the boundary conditions in Eqs. (9) through (12) gives the following linear equation system:

$$-H_n^{(1)}(k_mb)B_n+H_n^{(1)}(k_ib)D_n+H_n^{(2)}(k_ib)E_n=J_n(k_mb)A_n, \tag{13}$$

$$-\mu_mk_mH_n^{(1)\prime}(k_mb)B_n+\mu_ik_iH_n^{(1)\prime}(k_ib)D_n+\mu_ik_iH_n^{(2)\prime}(k_ib)E_n=\mu_mk_mJ_n'(k_mb)A_n, \tag{14}$$

$$H_n^{(1)}(k_ia)D_n+H_n^{(2)}(k_ia)E_n-J_n(k_fa)C_n=0, \tag{15}$$

$$\mu_ik_iH_n^{(1)\prime}(k_ia)D_n+\mu_ik_iH_n^{(2)\prime}(k_ia)E_n-\mu_fk_fJ_n'(k_fa)C_n=0. \tag{16}$$

III. ANALYTICAL SOLUTION

The linear equation system, Eqs. (13) through (16), can be solved to obtain closed-form expressions for various wave expansion coefficients.

First, Eqs. (15) and (16) can be used to solved for C_n and D_n in terms of E_n as

$$C_n = -\frac{4\hat{i}\mu_i}{\pi a \Delta} E_n, \tag{17}$$

$$D_n = -\frac{\overline{\Delta}}{\Delta} E_n, \tag{18}$$

where $\Delta = \Delta_1 + \hat{i}\Delta_2$

$$\Delta_1 = \mu_f k_f J_n'(k_f a) J_n(k_i a) - \mu_i k_i J_n(k_f a) J_n'(k_i a), \tag{19}$$

$$\Delta_2 = \mu_f k_f J_n'(k_f a) Y_n(k_i a) - \mu_i k_i J_n(k_f a) Y_n'(k_i a), \tag{20}$$

$Y_n(z)$ is a Bessel function of the second kind, and an overbar denotes the complex conjugate. In the process, the relation $H_n^{(1),(2)}(z) = J_n(z) \pm \hat{i} Y_n(z)$ and the following Wronskian of Bessel functions (Abramowitz and Stegun, 1965):

$$\mathcal{W}(J_n(z), Y_n(z)) = J_n(z) Y_n'(z) - J_n'(z) Y_n(z) = \frac{2}{\pi z}, \tag{21}$$

have been used.

Next, substituting Eqs. (17) and (18) into Eqs. (13) and (14) gives a pair of equations about B_n and E_n, which is subsequently solved. Afterwards, substituting the resulting expression for E_n into Eqs. (17) and (18) gives the final expressions for C_n and D_n in terms of A_n.

Finally, the complete solution can be written as

$$B_n = -\frac{\mu_m k_m M_1 J_n'(k_m b) - \mu_i k_i M_2 J_n(k_m b)}{\mu_m k_m M_1 H_n^{(1)\prime}(k_m b) - \mu_i k_i M_2 H_n^{(1)}(k_m b)} A_n, \tag{22}$$

$$C_n = -\frac{4\hat{i}}{\pi^2 a b} \frac{\mu_i \mu_m}{\mu_m k_m M_1 H_n^{(1)\prime}(k_m b) - \mu_i k_i M_2 H_n^{(1)}(k_m b)} A_n, \tag{23}$$

$$D_n = -\frac{\mu_m}{\pi b} \frac{\mu_f k_f J_n'(k_f a) H_n^{(2)}(k_i a) - \mu_i k_i J_n(k_f a) H_n^{(2)\prime}(k_i a)}{\mu_m k_m M_1 H_n^{(1)\prime}(k_m b) - \mu_i k_i M_2 H_n^{(1)}(k_m b)} A_n, \tag{24}$$

$$E_n = \frac{\mu_m}{\pi b} \frac{\mu_f k_f J_n'(k_f a) H_n^{(1)}(k_i a) - \mu_i k_i J_n(k_f a) H_n^{(1)\prime}(k_i a)}{\mu_m k_m M_1 H_n^{(1)\prime}(k_m b) - \mu_i k_i M_2 H_n^{(1)}(k_m b)} A_n, \tag{25}$$

where

$$M_1 = \Delta_2 J_n(k_i b) - \Delta_1 Y_n(k_i b), \tag{26}$$

$$M_2 = \Delta_2 J_n'(k_i b) - \Delta_1 Y_n'(k_i b). \tag{27}$$

IV. NUMERICAL RESULTS AND DISCUSSIONS

A. Parameters

Computations are performed for a ceramic-fiber-reinforced metal–matrix composite system whose constituents properties, after Yim and Williams (1995b), are listed in Table I. The fiber radius a is primarily used as a nondimensionalization length scale, whose value, which can be taken as $a = 10\,\mu$m when necessary, is never explicitly used in the computations. Unless otherwise noted, the outer radius of the interphase is assumed to be $b = 1.1a$.

TABLE I. Constituent material properties for a metal–matrix ceramic–fiber composite system (after Yim and Williams, 1995b).

Property	Matrix (AA520 aluminum)	Fiber (alumina, Al_2O_3)	Interphase (zirconia, ZrO_2)
Density (kg/m^3)	2600	3700	6300
Young's modulus (GPa)	66	360	97
Poisson's ratio	0.31	0.25	0.33
Lamé constant λ (GPa)	41	144	71
Shear modulus μ (GPa)	25	144	37
P wave speed (m/s)	5930	10800	4780
S wave speed (m/s)	3110	6240	2400

The incident wave is a planar wave of unit amplitude propagating along the $+x$ direction, which is expressible as (see, e.g., Pao and Mow, 1971)

$$\phi^{\text{inc}} = e^{\hat{i}k_m x} = e^{\hat{i}k_m r\cos\theta} = \sum_{n=-\infty}^{\infty} \hat{i}^n J_n(k_m r) e^{\hat{i}n\theta}. \tag{28}$$

Prior to computation, a simple criterion is established to truncate the infinite series that represent various wave fields in Eqs. (4) through (7). Consider the scattered wave in Eq. (5). Since $|H_n^{(1)}(k_m r)|$ monotonically decreases when either r

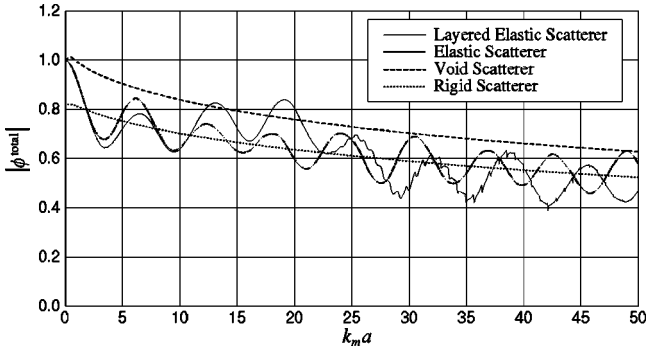


FIG. 2. Far-field forward response spectrum for fiber-interphase-matrix model (layered elastic scatterer), and three other models.

or n increases, and $r>a$, the series in Eq. (5) can be simply truncated at the Nth term when

$$|B_N H_N^{(1)}(k_m a)|<\varepsilon. \tag{29}$$

In the computations, $\varepsilon=10^{-8}$.

Response spectra are computed for the following arbitrarily chosen measurement points: the far-field forward and backward measuring points are $(x,y)=(50a,0)$ and $(-50a,0)$, respectively, and the near-field measuring points are $(\pm 5a,0)$.

B. Spectra as compared to other models

Figures 2 and 3 show the far-field and the near-field forward displacement response spectra; that is, the displacement amplitude $|\phi^{\text{total}}|$ versus the nondimensionalized frequency $k_m a$, for the fiber-interphase-matrix model. Also shown for comparison are the corresponding results for several simpler models: a circular elastic cylinder of radius a (the fiber-matrix model), a rigid circular cylinders of radius a, and a circular cylindrical void of radius a. For all these cases, the matrix medium remains the same. The fiber in the fiber-matrix model has the same properties as the fiber in the fiber-interphase-matrix model.

The analytical solution for the fiber-matrix model involves only the scattered wave in the matrix and the refracted wave in the fiber, as expressed in Eqs. (5) and (6), respectively. If the radius of the fiber is a and all other parameters remain the same, the solution can be found as (see, e.g., Pao and Mow, 1971)

$$B_n=-\frac{\mu_f k_f J_n'(k_f a)J_n(k_m a)-\mu_m k_m J_n(k_f a)J_n'(k_m a)}{\mu_f k_f J_n'(k_f a)H_n^{(1)}(k_m a)-\mu_m k_m J_n(k_f a)H_n^{(1)\prime}(k_m a)}A_n, \tag{30}$$

$$C_n=-\frac{2\hat{i}}{\pi a}\frac{\mu_m}{\mu_f k_f J_n'(k_f a)H_n^{(1)}(k_m a)-\mu_m k_m J_n(k_f a)H_n^{(1)\prime}(k_m a)}A_n. \tag{31}$$

Solutions for other two cases can be found in Pao and Mow (1971).

Two other models representing the degenerate cases, in which the interphase has the same material as either the matrix or the fiber, also could have been shown in the comparison. For brevity, however, it suffices to note that the results for these degenerate cases are identical to those for the elastic scatterer, with proper scaling of the geometry.

Figures 4 and 5 show the far-field and the near-field backward spectra for the displacement amplitude of the scattered waves; that is, $|\phi^s|$ versus $k_m a$.

In general, the total wave in the backward direction is highly oscillatory. This is due to the interference of the two waves traveling in opposite directions: the scattered wave travels in the backward direction, whereas the incident wave travels in the forward direction.

Recall the asymptotic expressions for Hankel functions (Abramowitz and Stegun, 1965) for large arguments z,

$$H_n^{(1)}(z)\to\sqrt{\frac{2}{\pi z}}\,e^{\hat{i}(z-[(1+2n)/4]\pi)}. \tag{32}$$

In the far field (when $k_m r$ is large), the total wave in the matrix can be written as

$$\phi^{\text{total}}=e^{\hat{i}k_m r\cos\theta}+\sum_{n=-\infty}^{\infty}B_n\sqrt{\frac{2}{\pi k_m r}}\,e^{\hat{i}(k_m r-[(1+2n)/4]\pi)}e^{\hat{i}n\theta}. \tag{33}$$

From Eq. (33), in the forward direction ($\theta=0$)

$$|\phi^{\text{total}}|_{\text{forward}}=|1+f_1(k_m)|, \tag{34}$$

and in the backward direction ($\theta=\pi$)

$$|\phi^{\text{total}}|_{\text{backward}}=|e^{-2\hat{i}k_m r}+f_2(k_m)|, \tag{35}$$

where

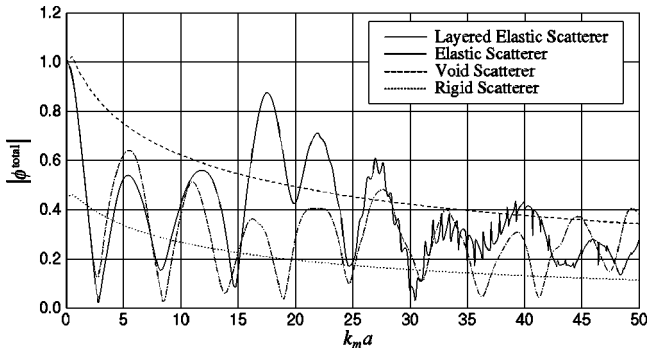


FIG. 3. Near-field forward response spectrum for fiber-interphase-matrix model (layered elastic scatterer), and three other models.

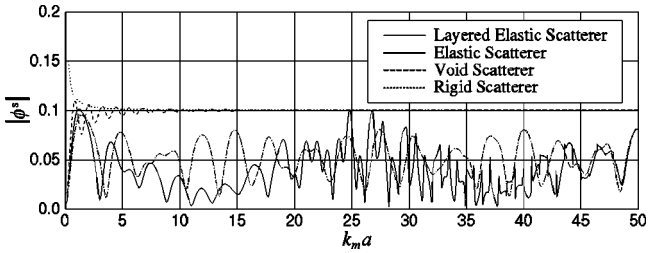


FIG. 4. Far-field backward response spectrum for fiber-interphase-matrix model (layered elastic scatterer), and three other models.

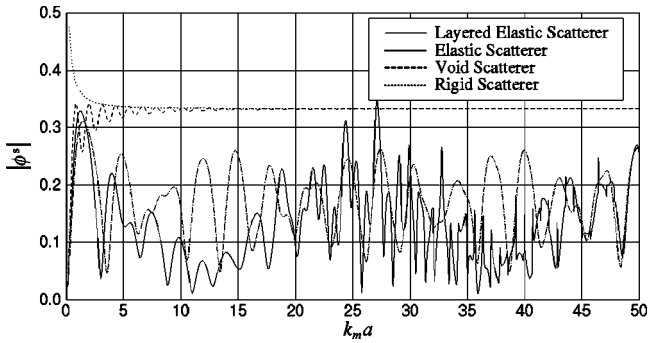


FIG. 5. Near-field backward response spectrum for fiber-interphase-matrix model (layered elastic scatterer), and three other models.

$$f_1(k_m) = \sqrt{\frac{2}{\pi k_m r}} \sum_{n=-\infty}^{\infty} B_n e^{-\hat{i}[(2n+1)/4]\pi},$$

$$f_2(k_m) = \sqrt{\frac{2}{\pi k_m r}} \sum_{n=-\infty}^{\infty} B_n e^{\hat{i}[(2n-1)/4]\pi}. \quad (36)$$

Therefore, for a fixed r in the frequency domain, high oscillation of $|\phi^{\text{total}}|_{\text{backward}}$ is attributed to the term $e^{-2\hat{i}k_m r}$ if the function $f_2(k_m)$ is not highly oscillatory. It follows that $|1 \pm f_2(k_m)|$ are the envelopes of the backward spectrum of the total displacement. For this reason, the scattered wave generally suffices for observing significant characteristics of the backward scattering.

Comparing the spectra for the elastic scatterer (fiber-matrix model) with those for the rigid and void scatterers in Figs. 2 through 5, it is observed that a spectrum for the elastic scatterer has an oscillatory structure whose ''period'' in the frequency domain appears to be approximately a constant. On the other hand, the spectra for both the rigid and void scatterers are almost monotonic, except in the low-frequency range. As compared with the rigid and void scatterer cases, the difference in the case of an elastic scatterer is that it has a refracted wave that resides within the scatterer. Therefore, the oscillatory structure can be attributed to the existence of this refracted wave.

Comparing the curves for the layered scatterer (fiber-interphase-matrix model) and the elastic scatterer in Figs. 2 through 5, the most notable difference is that there appears to be a second oscillatory structure in the spectra for the layered scatterer. At relatively low frequencies from $k_m a = 20$ to 32, which corresponds approximately to $k_i(b-a) = 2.2$ to 3.5, this secondary structure appears as a smooth oscillation that is superposed onto the main oscillatory structure. But, at relatively high frequencies from $k_m a = 42$ to 45, which corresponds approximately to $k_i(b-a) = 4.5$ to 5, the secondary structure turns into sharp spikes. At even higher frequencies $(k_m a > 45)$, the secondary structure tends to diminish.

C. Effects of interphase properties

Figure 6 shows the far-field spectrum maps when the ratio of the outer and inner radii b/a changes while all other parameters remain unchanged. In the figure, a displacement amplitude in the far-field measuring point is converted into a color according to the color scale shown at the right. Each line parallel to the abscissa represents a spectrum for the corresponding radii ratio. The displacement amplitude of the total wave is shown for the forward spectrum map, but only the displacement amplitude of the scattered wave is shown for the backward spectrum map.

Figure 6 provides a view of the overall trend of evolution of the second spectrum structure, which is the highly oscillatory component as observed in Figs. 2 through 5. It is observed from Fig. 6 that the interphase thickness, which is $b-a$, has a profound effect on the secondary spectrum structure. As the interphase thickness increases, the secondary spectrum structure starts to appear in the high-frequency regime, and gradually moves toward the lower frequency regime. The second oscillatory structure decreases its ''period'' in the frequency space as the frequency increases.

Figure 6 also shows the existence of the third spectrum structure that overall resembles the functional form of

$$k_m a(b/a - 1) = \text{constant}. \quad (37)$$

The most noticeable one of this spectrum structure is in the forward spectrum map as a series of orange blobs centered at $(k_m a, b/a) \approx (2,2)$, (7,1.25), (13,1.12), (19,1.09), (26,1.08), and (32,1.07). Correspondingly, it appears as blue streaks in the backward spectrum map.

Additionally, it is noted that the primary oscillatory structure that consistently appeared in Figs. 2 through 5 does not leave a traceable mark in this pair of spectrum maps. The primary oscillatory structure in those figures is dictated by the fiber, and theoretically would not change when any parameter of the interphase changes.

This suggests that the second and the third oscillatory structures are the results of the intertwining of two underlying oscillatory structures; one is dominated by the inner radius a and the other is dominated by the outer radius b. When their periods and strengths are not drastically different, the two intertwining oscillatory structures create a ''beat,'' resulting in two different oscillatory structure of the characteristics of

$$k_m b - k_m a = \text{constant}, \quad \text{and} \quad k_m b + k_m a = \text{constant}.$$

The former is the third oscillatory structure as identified in Eq. (37), and the latter is the second oscillatory structure.

Using a similar visualization technique to Fig. 6, Figs. 7 and 8 are the spectrum maps when the interphase shear modulus μ_i and mass density ρ_i change, respectively, while all other parameters remain unchanged.

In each spectrum map in Figs. 7 and 8, the upper and the lower halves of the spectrum maps appear dramatically different.

When $\log_{10}(\mu_i/\mu_m) > 0$ or when $\log_{10}(\rho_i/\rho_m) < 0$, the overall oscillatory structure of the spectrum is not altered in any significant way as the interphase parameter, either the shear modulus μ_i or the mass density ρ_i, is changed. Specifically, when the interphase is lighter than the matrix, the mass density of the interphase ρ_i has negligible effects on the spectrum; when the interphase is stiffer than the matrix, μ_i has only marginal effects on the spectrum.

However, when $\log_{10}(\mu_i/\mu_m) < 0$ or when $\log_{10}(\rho_i/\rho_m) > 0$, a series of curves overwhelms the oscillatory structure that dominates the other half of the spectrum

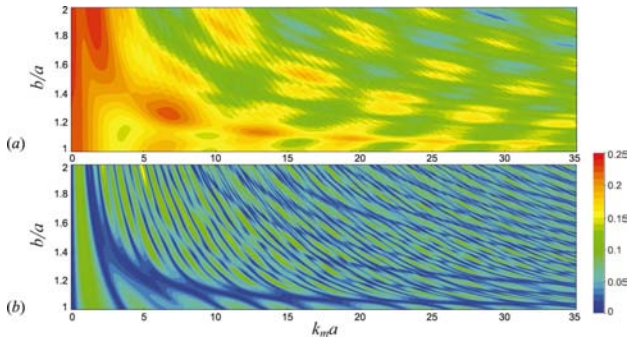


FIG. 6. Far-field spectrum maps when interphase thickness changes. (a) Forward spectrum map. (b) Backward spectrum map.

map. These curves would appear as oscillatory structures in the spectrum with the following two distinctive characteristics: (1) the oscillatory structure becomes sharp peaks in the low-frequency range, which appear in the spectrum map as large color change within a short frequency range; (2) there are secondary peaks or oscillatory structures that appear to be riding on top of the major peaks.

Since the geometry of the problem is unchanged in both Figs. 7 and 8, the explanation lies in the physical process. It is widely recognized that the oscillatory structure in the spectrum is due to various resonances in the scatterer (see, e.g., Pao and Mow, 1971). As indicated earlier, this paper was prompted by the discovery of a new mechanism (Liu *et al.*, 2000) for the formation of a sonic band gap in a cubically arranged array of scatterers. Such a band gap material or structure would have great potential for creating sound-proof structures.

Since the wavelength of audible sound wave in air ranges from a few millimeter to a few meters, the significance of this new discovery is making it possible to create acoustic filters or structural components in the size of a fraction of the wavelength.

Therefore, the focus here is on the lowest possible resonance peaks in these spectrum maps. Liu *et al.* (2000) had shown that for the spherical layered scatterer in their study the resonance at the lowest frequency corresponds to the

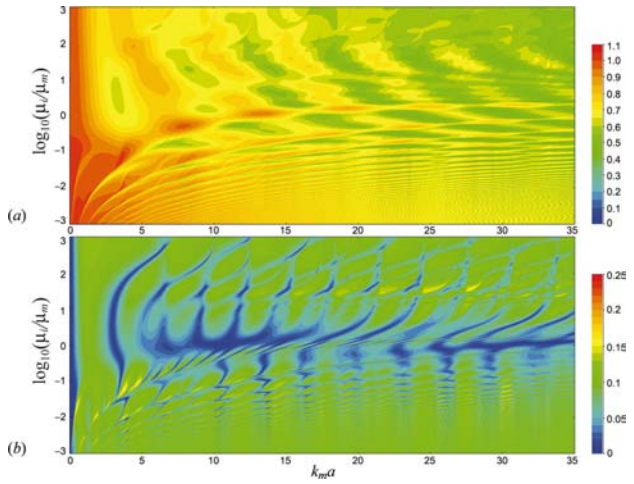


FIG. 7. Far-field spectrum maps when elastic modulus of interphase changes. (a) Forward spectrum map. (b) Backward spectrum map.

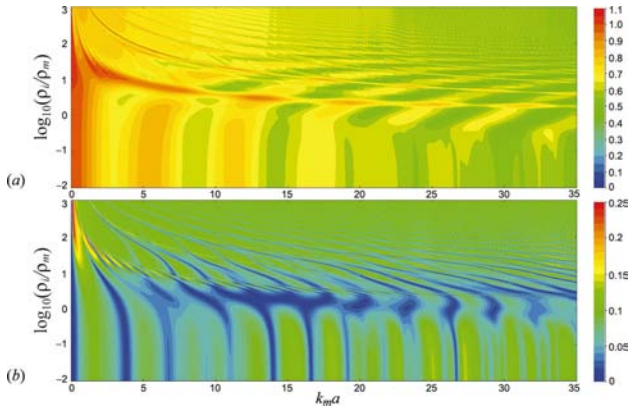


FIG. 8. Far-field spectrum maps when mass density of interphase changes. (a) Forward spectrum map. (b) Backward spectrum map.

rigid-body mode of the core, in which the core undergoes small deformation but notable rigid-body displacement. The core in their model corresponds to the fiber in the layered cylindrical scatterer in this paper.

There are resonance peaks in the extremely low-frequency range in each of Figs. 7 and 8. In Fig. 7, when the shear modulus of the interphase is much smaller than that of the matrix ($\log_{10}\mu_i/\mu_m < -1.5$), the first resonance peak appears in the frequency range between $k_m a = 0.1$ and 0.3. In Fig. 8, the resonance peak exists in the extremely low-frequency range between $k_m a = 0.05$ and 0.25 when the mass density of the interphase is much higher than that of the matrix ($\log_{10}\rho_i/\rho_m > 1.5$).

Recall that the wave in the fiber is expressed in Eq. (6). Because $J_0(0)=1$ and $J_n(0)=0$ for any $n\neq 0$ (see, e.g., Abramowitz and Stegun, 1965), the $n=0$ term in Eq. (6), $C_0 J_0(k_f r)$, represents the rigid-body mode of the fiber, and C_0 is a quantitative measure of the amplitude of the rigid body mode. Figure 9 shows the modulus of C_0, $|C_0|$, for the cases when $\log_{10}(\mu_i/\mu_m) = -2$ (solid curve) and $\log_{10}(\rho_i/\rho_m) = 2$ (dashed curve), respectively, with all other parameters unchanged.

Figure 9 confirms that first peaks of $|C_0| \sim k_m a$ curves for both cases match exactly the lowest peaks in Figs. 7 and 8 for the corresponding material properties ratios.

Furthermore, other peaks in the curve for $\log_{10}(\mu_i/\mu_m) = -2$ also match with subsequent peaks

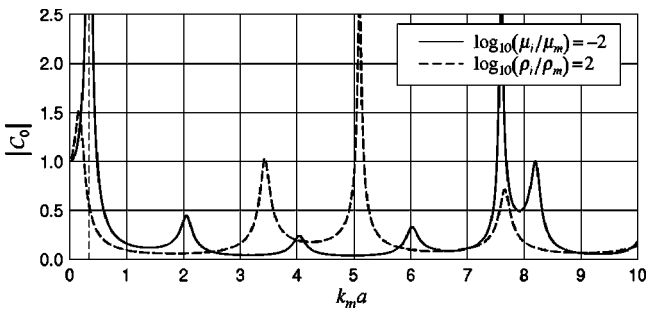


FIG. 9. Modulus of C_0 varies with frequency for cases of extremely compliant (solid curve) and extremely heavy (dashed curve) interphase materials. C_0 in Eq. (6) serves as measure of rigid-body mode in fiber. Vertical thin dashed line indicates lowest rigid-body mode frequency predicted by Eq. (40).

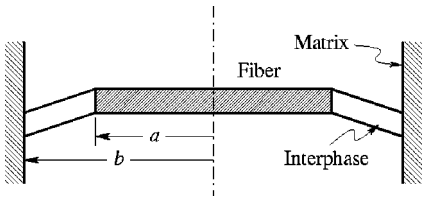


FIG. 10. Single degree-of-freedom model for predicting resonance frequency of rigid-body mode.

shown in Fig. 7. This indicates that the rigid-body mode of the fiber is the dominating mechanism for the main spectrum structure when the interphase is more compliant than the matrix.

On the other hand, the other peaks in the curve for $\log_{10}(\rho_i/\rho_m)=2$ only match some secondary streaks that appear as webs linking the major curves in the spectrum map in Fig. 7. This indicates that the rigid-body motion of the fiber is not the dominant mechanism when the interphase is heavier than the matrix. It is speculated that in such cases the dominating mechanism come from modes in the interphase.

For a rough estimate of the lowest resonance frequency of the rigid-body mode, a simple single degree-of-freedom mass–spring model is constructed. The model consists of a slice of fiber–interphase assembly attached to the surrounding matrix, which serves as the support. Figure 10 shows a deformed configuration of the model. It is assumed that the fiber is rigid and the interphase undergoes a radially linear displacement, having uniform shear stress and strain throughout. The equivalent spring constant is obtained from a force–deflection relation by imagining a force being applied at the center of the fiber, as

$$k_{\text{spring}}=\frac{2\pi a\mu_i}{b-a}. \tag{38}$$

The equivalent mass is obtained from the total kinetic energy by imagining the rigid fiber moving at a constant velocity and the interphase having a linear velocity distribution in the radial direction, as

$$m_{\text{eq}}=\pi a^2\rho_f+\tfrac{1}{6}\pi(b-a)(3a+b)\rho_i\,. \tag{39}$$

Thus, the frequency corresponding to the natural frequency of the mass–spring model, normalized by the fiber radius a, is

$$(k_m a)_n=\omega_n a\sqrt{\frac{\rho_m}{\mu_m}}=\sqrt{\frac{\dfrac{2a}{b-a}\dfrac{\mu_i}{\mu_m}}{\dfrac{\rho_f}{\rho_m}+\dfrac{1}{6}\left(\dfrac{b}{a}-1\right)\left(\dfrac{b}{a}+3\right)\dfrac{\rho_i}{\rho_m}}}. \tag{40}$$

Using this model, the predicted resonance frequency corresponding to the case of $\log_{10}(\mu_i/\mu_m)=-2$ is 0.355, which is satisfactorily close to the value of 0.34 as indicated in Figs. 7 and 9. In Fig. 9, the vertical thin dashed line indicates the resonance frequency predicted by Eq. (40). However, this model is only applicable to the cases when the interphase is extremely compliant.

V. CONCLUSIONS

The closed-form analytical solution for the scattering of a layered circular elastic cylinder has been presented. The following conclusions can be drawn from the numerical results.

(i) The presence of the interphase gives rise to the second geometric parameter for the problem, and accordingly, the spectrum in general consists of two oscillatory structures whose periodicities are dominated by $k_m(b+a)$ and $k_m(b-a)$.

(ii) The presence of the interphase also gives rise to the possible rigid-body mode of the fiber. The lowest resonance frequency could be extremely low. Rigid-body mode occurs in two extreme cases: either the interface is extremely compliant, or the fiber is extremely heavy.

(iii) The rigid-body mode of the fiber is the dominating mechanism for the main spectrum structure when the interphase is more compliant than the matrix.

(iv) A simple single degree-of-freedom model gives a satisfactorily close prediction for the lowest frequency of the rigid-body mode when the interphase is extremely compliant.

Most of these conclusions are not quantitative, but point out some interesting phenomena in single scatterer problems in general, and the fiber-interphase-matrix model in particular. The strong resonance in the extremely low-frequency range may lead to new design principles for sonic band gap materials.

ACKNOWLEDGMENTS

The author is grateful to Professor James H. Williams, Jr., of the Department of Mechanical Engineering of Massachusetts Institute of Technology, for his guidance and encouragement in the early part of this work. The author is also grateful to an anonymous reviewer for his suggestion of constructing the single degree-of-freedom model for predicting the rigid-body mode frequency.

Abramowitz, M., and Stegun, I. (**1965**). *Handbook of Mathematical Functions* (Dover, New York).

Addison, R. C., Jr., and Sinclair, A. N. (**1992**). "Calculated and measured ultrasonic response of an elastic cylinder embedded in an elastic medium," in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York), Vol. 11A, pp. 105–111.

Bogen, S. D., and Hinders, M. K. (**1993**). "Dynamics stress concentration in fiber-reinforced composites with interface layers," J. Compos. Mater. **27**, 1272–1311.

Bogen, S. D., and Hinders, M. K. (**1994**). *Interface Effects in Elastic Wave Scattering* (Springer, Berlin).

Cai, L.-W. (**1998**). "Full-scale simulation of multiple scattering of elastic waves in fiber reinforced composites," Sc.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Cai, L.-W., and Williams, J. H., Jr. (**1999a**). "NDE via stop band formation in fiber reinforced composites having square fiber arrangements," Ultrasonics **37**, 483–492.

Cai, L.-W., and Williams, J. H., Jr. (**1999b**). "Full-scale simulations of elastic waves scattering in fiber reinforced composites," Ultrasonics **37**, 463–482.

Cai, L.-W., and Williams, J. H., Jr. (**1999c**). ''Large-scale multiple scattering problems,'' Ultrasonics **37**, 453–462.
Chu, Y. C., and Rokhlin, I. (**1992**). ''Determination of macro- and micromechanical and interfacial elastic properties of composites from ultrasonic data,'' J. Acoust. Soc. Am. **92**, 920–931.
Huang, W., Brisuda, S., and Rokhlin, S. I. (**1995**). ''Ultrasonic wave scattering from fiber-matrix interphases,'' J. Acoust. Soc. Am. **97**, 807–817.
Liu, Z., Zhang, X., Mao, Y., Zhu, Y. Y., Yang, Z., Chan, C. T., and Sheng, P. (**2000**). ''Locally resonant sonic materials,'' Science **289**, 1734.
Mal, A. K., and Yang, R.-B. (**1994**). ''The influence of the fiber-matrix interfacial layer properties on wave characteristics in metal matrix composites,'' in *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York), Vol. 13B, pp. 1453–1460.
Pao, Y.-H., and Mow, C.-C. (**1971**). *Diffraction of Elastic Waves and Dynamic Stress Concentrations* (Crane Russak, New York).
Shindo, Y., and Niwa, N. (**1996**). ''Scattering of antiplane shear wave in a fiber-reinforced composite medium with interfacial layers,'' Acta Mech. **117**, 181–190.
Sinclair, A. N., and Addison, R. C., Jr. (**1993**). ''Acoustic diffraction spectrum of a SiC fiber in a solid elastic medium,'' J. Acoust. Soc. Am. **94**, 1126–1135.
Yim, H., and Williams, J. H., Jr. (**1995a**). ''Formulation and its energy balance verification for ultrasonic nondestructive characterization of a single fiber composite interphase,'' Ultrasonics **33**, 377–387.
Yim, H., and Williams, J. H., Jr. (**1995b**). ''Database generation and parametric study for ultrasonic nondestructive characterization of a single fiber composite interphase,'' Ultrasonics **33**, 389–401.

Ultrasound diffusion for crack depth determination in concrete

S. K. Ramamoorthy, Y. Kane, and J. A. Turner[a)]
Department of Engineering Mechanics, W317.4 Nebraska Hall, University of Nebraska-Lincoln, Lincoln, Nebraska 68588-0526

(Received 2 July 2003; revised 15 November 2003; accepted 24 November 2003)

The determination of the depth of surface-breaking cracks in concrete specimens using an ultrasound diffusion technique is discussed. Experiments were carried out on precracked concrete specimens of varying crack depths (0%–40% of the specimen thickness). Contact transducers were placed at the specimen surface with source and receiver separated by the crack. Tone-burst excitations over a frequency range of 400–600 kHz were used. At these frequencies, ultrasound is scattered considerably by the heterogeneities in the concrete. In the limit of many scattering events, the evolution of energy may be modeled as a diffusion process. The arrival of the peak diffuse energy at the receiver is delayed due to the presence of crack. This delay is the prime indicator used for determining crack depth. Numerical and analytical analyses were also used for comparison. These results are in basic agreement with the experiments. In addition, these analyses are used to study the limits of this technique. In particular, it is shown that this technique is applicable to cracks greater than the scattering mean-free path, which is estimated at about 1 cm for these specimens. Aspects of practical implementation are also discussed.
[DOI: 10.1121/1.1642625]

PACS numbers: 43.20.Gp, 43.35.Cg, 43.35.Zc [SFW] Pages: 523–529

I. INTRODUCTION

Concrete structures are continually subject to effects that degrade their structural integrity. Near-surface damage in concrete may take the form of distributed microcracks or distinct, large cracks that extend to the surface of the structure. Damage from microcracking develops at length scales that are smaller than the size of the majority of the aggregate.[1] However, large cracks can result from mechanical overloading, freeze–thaw cycling, by drying shrinkage, or from the coalescence of many microcracks. Several recent research articles report efforts to determine nondestructively the depth of surface-breaking cracks in concrete using elastic stress wave time-of-flight (wave velocity) techniques.[2–4] In these papers, the crack depth is determined if the velocity of wave propagation in the intact concrete is known and a particular wave path and wave pulse–crack interaction are assumed. The frequency range for these experiments varied from 0 to 100 kHz. A frequency of 100 kHz corresponds to a wavelength on the order of a few centimeters in concrete. In the impact–echo method,[2] the user must interpret the recorded echo signals. Quality interpretation requires a user experienced in waveform recognition. Other researchers use frequencies in the range of 0 to 70 kHz.[5,6] The wave source is due to the impact of a steel ball on the specimen surface, as in the impact–echo method. Surface wave methods have also been used to characterize the microstructural properties of concrete such as grain size, porosity, or microcrack distribution.[7] The penetration depth of the surface waves is on the order of one wavelength. Hence, methods that utilize surface waves are, in general, insensitive to deep cracks. To detect deep cracks by surface wave methods, low-frequency waves must be used. These can be difficult to generate with sufficient energy without damaging the surface. Ultrasonic guided waves (Lamb waves) have also been used to detect damage in concrete beams and to detect delaminations between steel bars and concrete.[8] However, guided wave testing is complicated by the presence of the many available modes and by the difficulties associated with deep slabs. It is also difficult to generate a low-frequency guided wave using commercial transducers.

The response of a heterogeneous elastic body to a propagating wave is a function of the wavelength in comparison with the length scale of the heterogeneities. When wavelengths on the order of the microstructure length scale are used, a large amount of scattering occurs. The scattering process causes the elastic wave energy to propagate in directions which do not coincide with the incident wave. The propagating wave is attenuated due to these scattering losses. The scattering process is generally energy conserving—the scattered energy is not lost, but is no longer in phase with the exciting wave. Additional attenuation is caused by true dissipative mechanisms. Thus, the material response becomes a combination of coherent and diffuse (incoherent) energy. If the initial energy scatters many times during its path from source to receiver, it is expected to behave as a diffusion process. Previous one-dimensional diffusion experiments in concrete have shown that the ultrasonic diffusivity and dissipation may be extracted from such measurements.[9] In many cases, the evolution of the diffuse energy occurs on a much longer time scale than the time necessary for the coherent wave to propagate across the specimen.

In this article, a new technique is discussed to determine the depth of surface cracks in concrete by exploiting the multiple scattering properties of the concrete. In the next section, the concrete specimens are discussed. In Sec. III, the

a)Author to whom correspondence should be addressed. Electronic mail: jaturner@unl.edu

TABLE I. Specimen details.

Specimen ID	Dimensions (cm)	Crack depth (cm)	% Crack depth to slab depth
Slab 1	60.96×60.96×20.32	0	0
Slab 2	60.96×60.96×20.32	1.27	6.25
Slab 3	60.96×60.96×20.32	3.81	18.75
Slab 4	60.96×60.96×20.32	7.62	37.50

results for the uncracked slab are presented. The one-dimensional diffusion results[9] are extended to two dimensions representative of the slab geometry considered here. The ultrasonic diffusivity and dissipation were determined from experiments by comparing the experimental results with an analytical solution of the two-dimensional diffusion process. These parameters may also be extracted from experiments on cracked slabs if a location away from the major crack is used. For all measurements, the diffuse energy field was determined by low-pass filtering the square of the transducer response. In addition, these results were compared with the solution obtained numerically from the two-dimensional diffusion equation. In Sec. IV, the analysis for the slabs with a major crack is presented. The experiments were conducted on three cracked specimens. The time shift of the peak amplitude of the diffuse energy can be used to determine the depth of the surface-breaking crack. Based on dimensional analysis, the time shift is expected to be proportional to the square of the crack depth divided by the diffusivity. This scaling with crack depth has been confirmed. The experimental values were in agreement with the numerical results. The potential advantages of this technique over other methods include the ease of excitation at high frequencies and the sensitivity to deep cracks. Also, it is not necessary to determine the wave velocity and other wave propagation parameters (wave path, interaction with geometry of the slab). The limits of this technique are its insensitivity to shallow cracks and possibly low signal-to-noise ratio, although sufficient repetition averaging can be used to overcome the latter.

II. SPECIMEN PREPARATION

Four specimens were cast with 47 BD bridge deck slab concrete having a compressive strength of 35 MPa with varying notch depth (0%–40%) of the thickness. The notches were created by placing a steel plate of 3-mm thickness in the mold before casting it with concrete. Cast oil was applied to both sides of the steel plate to facilitate easy removal. After the initial setting time, the steel plates were removed from the mold and the specimens were covered with burlap. After 24 h, the specimens were demolded and cured in a fog room for the next 28 days in order to reduce the likelihood of shrinkage cracks. Specific details of the specimens are shown in Table I.

III. UNCRACKED SLABS

Previously, the diffusion of ultrasound in concrete cylinder specimens was modeled as a one-dimensional process which does not apply to the slab geometry of interest here.[9] To recover the diffusivity and dissipation, the diffuse energy evolution was modeled as a two-dimensional diffusion process. The analytical solution was compared with the experimental solution to determine the scattering properties of the concrete. In this section, the extension of the one-dimensional results to the two-dimensional diffusion problem relevant for these slabs is discussed in terms of theory, numerical modeling, and experiments.

A. Analytical solution

The diffusion of ultrasonic energy through a concrete slab with no crack is modeled by the two-dimensional diffusion equation with dissipation. The equation is given by

$$D\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\langle E(x,y,t)\rangle-\frac{\partial}{\partial t}\langle E(x,y,t)\rangle-\sigma\langle E(x,y,t)\rangle = f(x,y,t), \tag{1}$$

with forcing condition

$$f(x,y,t)=E_0\,\delta(x-x_0)\,\delta(y-y_0)\,\delta(t-t_0), \tag{2}$$

where $E(x,y,t)$ is the ultrasonic spectral density (energy per area, per frequency bandwidth), D is the ultrasonic diffusivity with dimension length squared per time, σ is the dissipation with dimension inverse time, and E_0 is the initial energy which is deposited at time $t=0$. Equation (1) describes the evolution of the ensemble average energy density for an infinite medium. For the case considered here, Neumann boundary conditions (zero flux across the boundaries) are enforced at the specimen boundaries. The series solution of Eq. (1) for the source located at x_0 and y_0 is given by

$$\begin{aligned}E(x,y,t)=\Bigg\{&1+\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}4\cos\left(\frac{n\pi x_0}{l}\right)\cos\left(\frac{m\pi y_0}{p}\right)\\&\times\cos\left(\frac{n\pi x}{l}\right)\cos\left(\frac{m\pi y}{p}\right)e^{-D[(n\pi/l)^2+(m\pi/p)^2]t}\\&+\sum_{n=1}^{\infty}2\cos\left(\frac{n\pi x_0}{l}\right)\cos\left(\frac{n\pi x}{l}\right)e^{-D[(n\pi/l)^2]t}\\&+\sum_{m=1}^{\infty}2\cos\left(\frac{m\pi y_0}{p}\right)\cos\left(\frac{m\pi y}{p}\right)e^{-D[(m\pi/p)^2]t}\Bigg\}\\&\times E_0e^{-\sigma t},\end{aligned} \tag{3}$$

where l and p are the lateral dimensions of the slab.

Figure 1 shows example energy curves calculated using Eq. (3) for different combinations of ultrasonic diffusivity and dissipation values using $l=0.6$ m and $p=0.2$ m. The vertical positions of the source and receiver are identical ($y=y_0=0.2$ m), while the horizontal positions are equidistant from the center of the slab and separated by 6 cm ($x_0=0.27$ m and $x=0.33$ m) Steeper curves correspond to higher dissipation. The peak arrival time of the diffuse energy is an important parameter in determining the depth of the surface-breaking crack. Figure 1 shows that this peak arrival time decreases for higher values of dissipation.

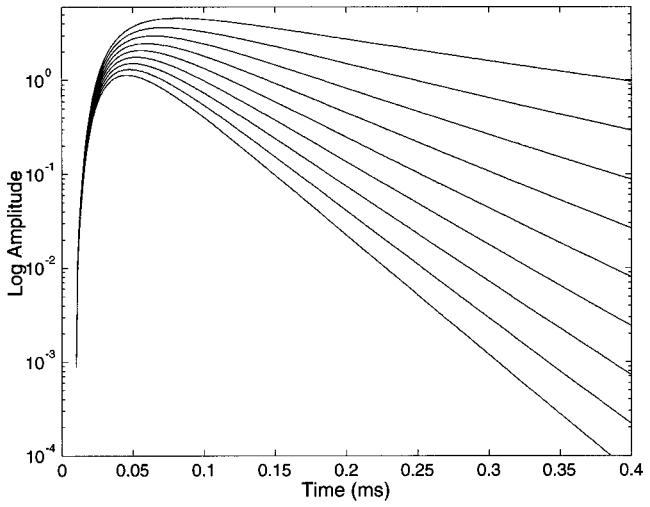


FIG. 1. Analytical solution, Eq. (3), for diffusivity of 12 m^2/s and dissipation varying from 3000 to 27 000 1/s assuming slab dimensions $l=0.6$ m and $p=0.2$ m. The vertical positions of the source and receiver are identical ($y=y_0=0.2$ m), while the horizontal positions are equidistant from the center of the slab and separated by 6 cm ($x_0=0.27$ m and $x=0.33$ m).

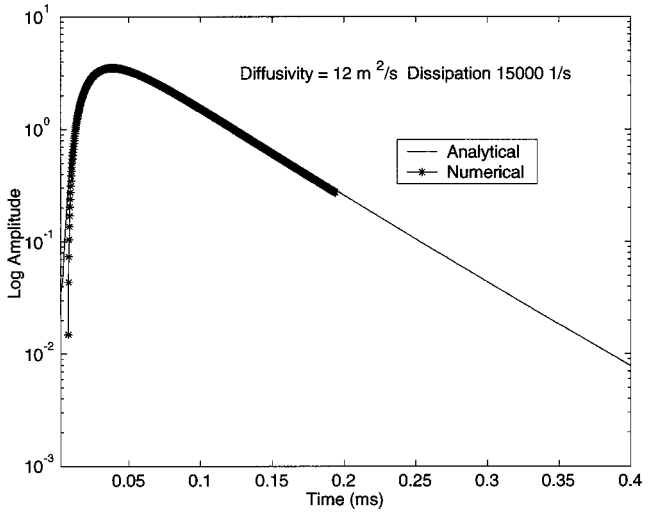


FIG. 2. Diffuse energy for slab 1 (uncracked) by analytical and numerical methods for a separation distance of 6 cm.

B. Numerical solution

A finite-element code was developed for solution of the diffusion equation given by Eq. (1) with forcing condition given by Eq. (2). The cross-sectional details of the experimental specimens given in Table I were used for creating a rectangular domain for this model. The rectangular domain was discretized and meshed with bilinear rectangular elements, and the boundary of the domain was modeled with the appropriate Neumann boundary conditions. By using the divergence theorem, the weak formulation of the diffusion Eq. (1) is given by

$$\int \left[\nu\left(\frac{\partial E}{\partial t} - \sigma E - E_0 \right) + D\frac{\partial \nu}{\partial x}\frac{\partial E}{\partial x} + D\frac{\partial \nu}{\partial y}\frac{\partial E}{\partial y} \right] dx\, dy = 0, \quad (4)$$

where ν is the weight function, chosen from the basis function $\psi_j^e(x,y)$ used for calculating the approximate energy, E, and E_0 is the initial energy deposited. The finite-element approximation is given by

$$E(x,y,t) \approx \sum_{j=1}^{n} E_j^e(t)\psi_j^e(x,y), \quad (5)$$

where E_j is the value of $E(x,y,t)$ at a point (x_j,y_j) at time t and $\psi_j^e(x,y)$ is the basis function. In matrix form, the transient energy equation at the elemental level is given by

$$[M^e]\{E^e\} + [K^e]\{E^e\} = \{f^e\}. \quad (6)$$

A forward difference scheme was used to solve Eq. (6). The time step Δt for the numerical analysis was chosen such that the stability criterion given by

$$\Delta t < \Delta t_{cr} = \frac{2}{\lambda_{\max}}, \quad (7)$$

is satisfied. In Eq. (7), $\lambda_{\max}$ is the largest eigenvalue of $[M^{-1}][K]$, where $[M]$ and $[K]$ are the global mass and stiffness matrices, respectively.

The typical time step used in the numerical analysis ranged from 0.2 to 2 μs. To check for convergence of the code, the analysis was carried out by varying the number of elements in the rectangular domain. It was found that the numerical solution compared well with the analytical solution (with a relative error of a few microseconds for the peak arrival time) for 1200 elements. Increasing the number of elements to 2400 increased the computation time considerably and did not significantly affect the peak arrival time. Since the main measure of crack depth is the peak arrival time, 1200 elements was judged sufficient for resolving this peak time. Figure 2 shows the comparison of the diffuse energy obtained by analytical and numerical methods for $D=12$ m^2/s and $\sigma=15\,000$ 1/s. The peak arrival time depends on the diffusion parameters D and σ as well as the source–receiver separation distance. This dependence is illustrated in Fig. 3. The peak arrival time, calculated numerically, is plotted versus source–receiver separation distance for various combinations of D and σ. The peak arrival time varies from tens of μs to hundreds of μs. It can also be seen that the

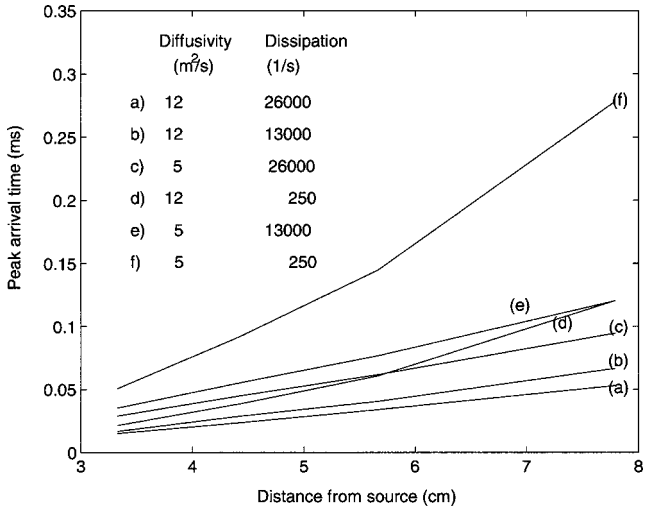


FIG. 3. Numerical results for peak arrival time as a function of source–receiver separation and ultrasonic diffusion parameters D and σ.

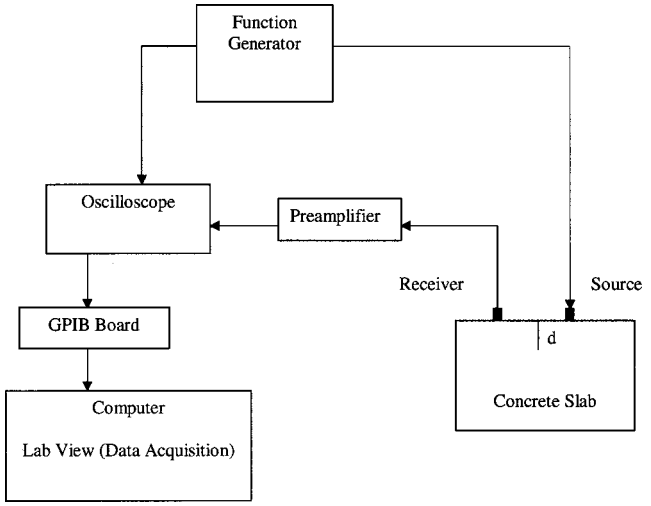


FIG. 4. Schematic layout of experimental setup.

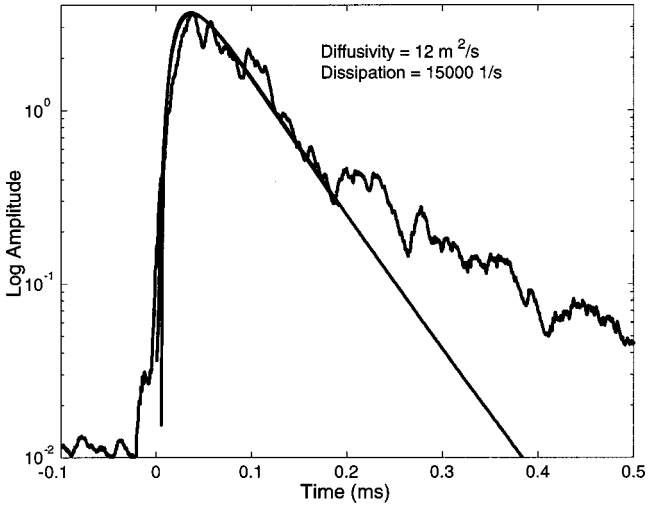


FIG. 5. Comparison of experimental data with analytical and numerical results to recover the scattering coefficients.

separation distance is more critical, in terms of peak arrival time, for certain combinations of D and σ.

C. Experiments

Experiments were first conducted on slab 1 (without a crack) for comparison with the analytical solution given by Eq. (1). The schematic layout of the ultrasound diffusion experiment is shown in Fig. 4. The source transducer, a 25.4-mm Panametrics contact transducer with a central frequency of 500 kHz, was coupled to the top surface of the slab. The receiving transducer, a Valpey-Fisher point-like pin transducer (2.5-mm diameter) was located at a separation distance r from the center of the source. Both the source and receiver were coupled to the specimen with a viscous couplant. The diffuse energy was obtained for different separation distances. The transducers were placed away from the edges to avoid edge effects. The source transducer was excited with a 5-cycle tone burst over a frequency range of 400–600 kHz using an arbitrary waveform generator. The output from the receiving transducer was preamplified and then digitized by an oscilloscope. To improve the signal-to-noise ratio, repetition averaging was used (500 averages). The oscilloscope was connected to a personal computer through a GPIB board and a LABVIEW program was used to acquire the average waveform. The square of the waveform data was then filtered using a low-pass filter. The result is a quantity proportional to the diffuse energy. The diffusivity and dissipation of the concrete slab were obtained by comparing the energy data with the analytical and numerical results.

Figure 5 shows the comparison of the experimental diffuse energy with the analytical and numerical results for an input frequency of 500 kHz at a source–receiver separation distance of 6 cm. The experimental results behave according to the proposed model only up until about 0.2 ms. After that time, the decay is no longer exponential. This type of behavior was often observed and is presumed due to the variation of dissipation within the frequency band. At these frequencies, the change of dissipation can be dramatic.[9] The analytical results were computed using 50 terms in Eq. (3). No significant change occurs in the solution if the number of terms is increased beyond 50. The diffusivity and the dissipation in Eq. (3) were adjusted until the analytical and experimental curves were best matched. Since the initial energy deposited, E_0, is not known, the amplitude of the analytical solution was scaled to match with the experimental data. The scattering parameters obtained by comparing the experimental data with the analytical solution for different separation distances were in agreement with the data obtained at the minimum separation distance between the source and receiver. Experiments were also carried out for the other slabs (slab 2, slab 3, and slab 4) at a location away from the surface-breaking crack. A comparison of the diffuse energy data for the different slabs confirms the assumption that the scattering properties of the concrete slab can be determined by measuring the diffuse energy field at locations away from major cracks and boundaries.

IV. CRACKED SLABS

The cracked slabs have surface breaking cracks of varying depth (slab 2, slab 3, and slab 4). The scattering properties recovered from the uncracked slab were used to determine the predicted peak arrival time in the numerical analysis for different crack depths. The lag time, defined as the difference between the peak arrival time of the diffuse energy in the cracked and uncracked regions, is the measure used here to infer the crack depth. Comparison between experimental and numerical lag times was used to evaluate this technique.

A. Analytical solution

For the cracked slab case, the governing equation remains the same. However, in addition to the Neumann boundary conditions at the slab boundaries, an additional boundary condition is imposed, i.e., no flux across the crack. Because an analytical solution for the problem of a slab with a crack is not yet available, the numerical solution was the primary measure for comparison with the experiments on cracked slabs.

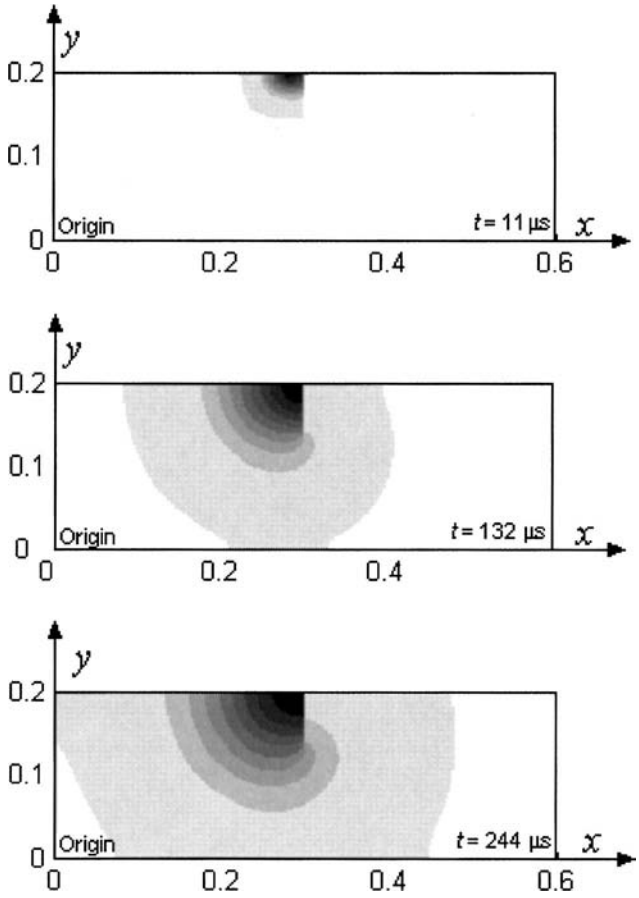


FIG. 6. Diffuse energy field for a crack depth of 30% at 11, 132, and 244 μs ($D=12\ \mathrm{m}^2/\mathrm{s}$, $\sigma=15\,000$ 1/s).

B. Numerical solution

For the cracked slabs the mesh discretization, element type, and number of elements were similar to those used for the uncracked slab. The crack was created at a given location by adding another set of nodes at the same location as existing nodes to the required crack depth. The element connectivity of the nodes at the crack location was changed such that the zero-flux condition across the crack is satisfied. Figure 6 shows the full diffuse energy field at times 11, 132, and 244 μs, respectively, for a crack depth of 30% of the slab depth ($D=12\ \mathrm{m}^2/\mathrm{s}, \sigma=15\,000$ 1/s). The diffusion process and the delay in the arrival of the diffuse energy across the crack is clearly depicted. The analysis was carried out for different combinations of diffusivity and dissipation. The results clearly show the delay in the peak arrival time of the diffuse energy across the crack when compared with the diffuse energy for the same source–receiver separation distance at the uncracked region. Figure 7 shows the lag time for different crack depths and for different source–receiver separation distances for select values of diffusivity and dissipation. The trend is similar for all separation distances, suggesting that this distance is not critical for practical implementation. It is also clear that the lag time is a function of separation distance between the source and transducer as expected.

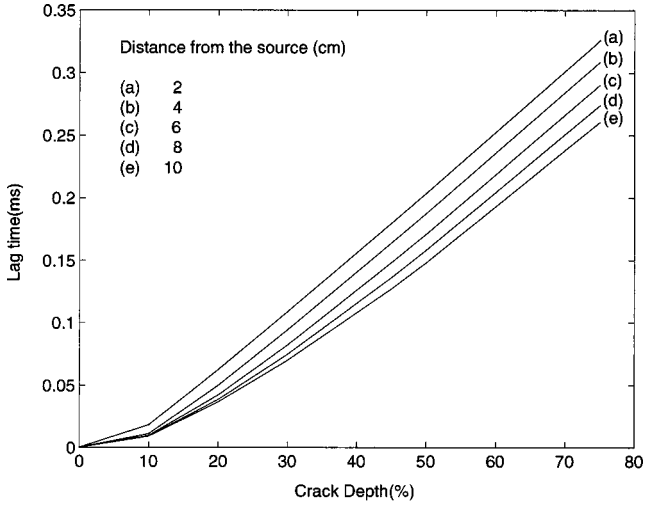


FIG. 7. Numerical results of lag time vs crack depth for different source–receiver separation distances ($D=12\ \mathrm{m}^2/\mathrm{s}$ and $\sigma=15\,000$ 1/s).

C. Experiments

The experiments on the cracked slabs were essentially identical to those for the uncracked slab. The source transducer was placed on one side of the crack and the receiver was placed on the other side of the crack as shown in Fig. 4. Figure 8 shows the waveform data obtained by the receiver for which the source and receiver were placed in an uncracked region. Figure 9 shows the waveform data for the source–receiver pair placed across the crack. Comparison of Figs. 8 and 9 shows the delay in the diffusion process clearly. There is a delay in the arrival of the peak diffuse energy when compared with the arrival of the diffuse energy obtained by receiver placed at the uncracked region of the slab. The delay in diffusion is due to the absence of flux transfer across the crack. This delay in arrival of the energy is the prime factor in assessing the depth of the surface opening crack.

The diffuse energy was obtained from the waveform data by squaring the transducer signal and filtering with a low-pass filter. The cutoff frequency for the low-pass filter

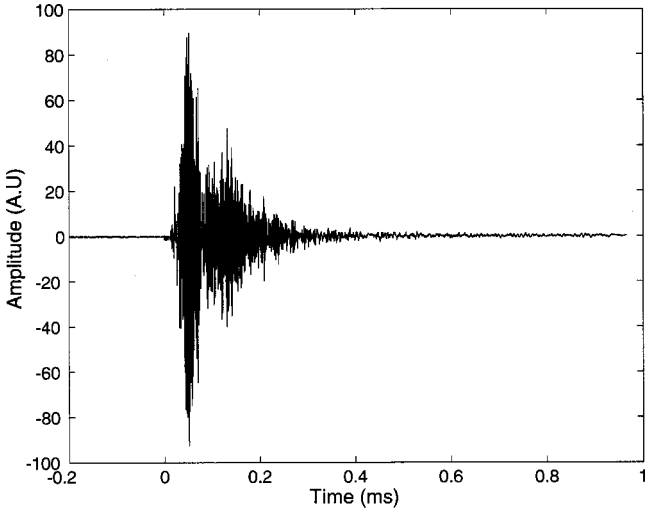


FIG. 8. Typical waveform data in an uncracked zone for an input frequency of 500 kHz and a source–receiver separation distance of 6 cm.

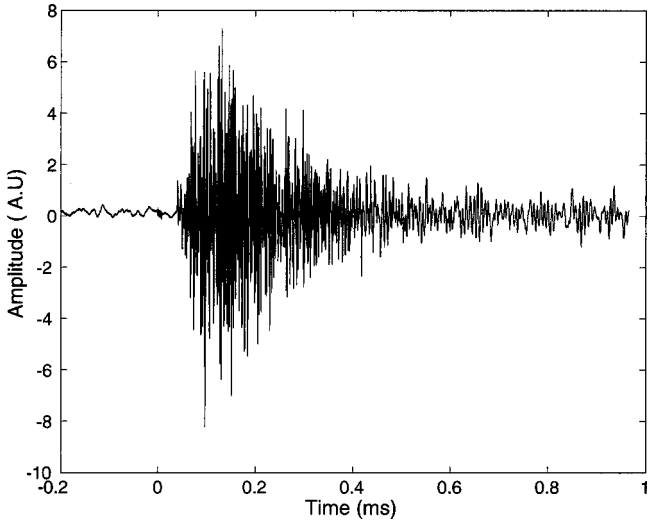


FIG. 9. Typical waveform data across the crack for an input frequency of 500 kHz and a source–receiver separation distance of 6 cm.

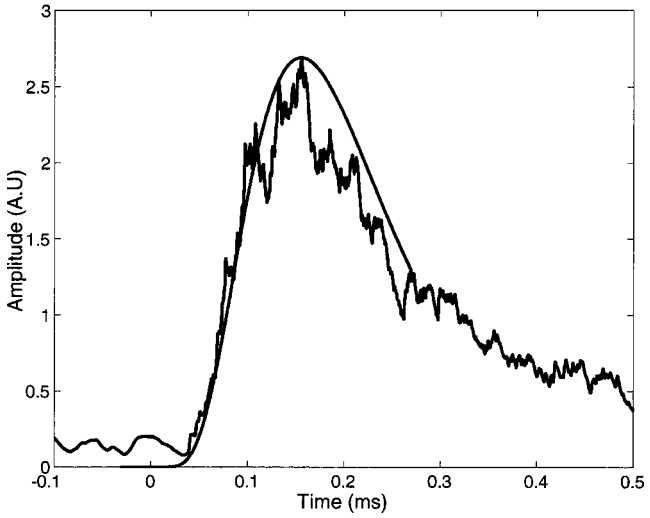


FIG. 11. Comparison of experimental diffuse energy with numerical results for slab 4 (37.5% crack depth).

was selected by considering the stability of the energy data with respect to cutoff frequency and by considering the fluctuations in the energy data. The results were scaled such that the initial rise of the experimental and numerical data matched. Figures 10 and 11 show the comparison of the experimental diffuse energy with the numerical data for slabs 3 and 4, respectively, with a separation distance of 6 cm between the source and the receiver for a cutoff frequency of 10 kHz. By comparing the peak arrival time of the waveform across the crack with the peak arrival time of the data obtained at the uncracked region, the lag time for different crack depths was determined. This lag time is due to the presence of the crack. Figure 12 shows the comparison of lag time obtained by numerical analysis and by experiments for different crack depths. The error bars for the experimental lag time were determined by low-pass filtering the waveform data for three different cutoff frequencies (4, 10, and 25 kHz). The standard deviation and average values of the three peak arrival times were used to calculate the error bars. For deeper cracks, the peak arrival time was more sensitive to the cutoff frequency than it was for shallow cracks. The agreement between the experimental results and the numerical results is very good. The inversion of experimental results to determine crack depth can occur if the diffusivity and dissipation are determined in a sample region away from boundaries and large cracks.

Trends observed in numerous numerical simulations show that the lag time can always be fit by a second-order polynomial function of crack depth.[10] The quadratic term is inversely proportional to diffusivity as expected, and the linear term is inversely proportional to the dissipation. This curve must pass through the origin since there is no lag time for a zero-depth crack. Thus, there is not a constant term in this polynomial. The experimental data show that this predicted dependence does indeed hold, at least for the values of D and σ relevant for the concrete samples examined here.

V. CONCLUSIONS

In this article, it has been shown that in light of many scattering events the evolution of ultrasonic energy in a con-

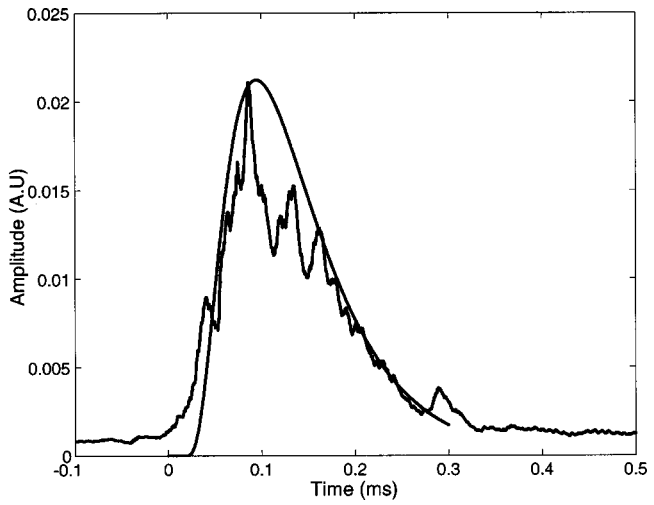


FIG. 10. Comparison of experimental diffuse energy data with numerical result for slab 3 (18.75% crack depth).

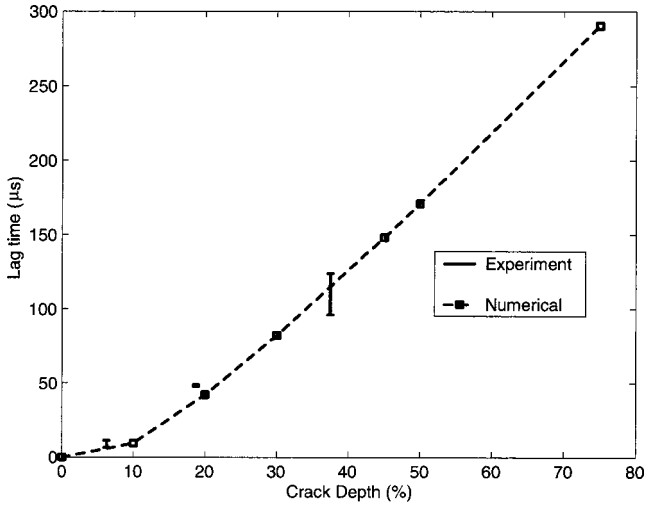


FIG. 12. Lag time vs crack depth for a source–receiver separation distance of 6 cm. The solid line is a quadratic curve fit to the numerical data and the error bars correspond to the experimental results.

crete slab can be modeled as a two-dimensional diffusion process. The ultrasonic diffusivity and dissipation were determined by comparing the experimental results with the analytical solution for a concrete slab without a crack. The results also compared well with the experimental results obtained in regions of the specimens that are not near boundaries or large cracks. The lag time was obtained by comparing the peak arrival time of the diffuse energy obtained at two locations, one in an uncracked region and the other across the crack. The experimental and numerical results correlated well for deeper cracks. The lag time increases considerably with increase in crack depth. The depth of the surface-breaking crack in a given concrete slab can be determined in two steps. In the first step, the scattering properties are determined from measurements in the uncracked region of the sample. In the second step, using the recovered scattering properties, the relation between the lag time and crack depth is obtained by numerical analysis. Once the relation is established, the depth of the surface-breaking crack is determined by using the experimental lag time. It should be noted that the results presented here are for notches, for which the two surfaces were not in contact. Results for real cracks with surfaces in contact are the focus of future research. The potential for characterizing nonstraight cracks and fluid-filled cracks is also under investigation.

ACKNOWLEDGMENTS

The support of the National Science Foundation (Grant Nos. CMS-9978707 and INT-0089548) and the National Bridge Research Organization (NaBRO) at the University of Nebraska-Lincoln is gratefully acknowledged by the authors.

[1] S. Popovics and J. S. Popovics, ''Ultrasonic testing to determine water–cement ratio for freshly mixed concrete,'' Am. Stand. Test. Mater. **20**, 262–268 (1998).
[2] M. Sansalone, J. M. Lin, and W. B. Streett, ''Determining the depth of surface-opening cracks using impact-generated stress waves and time-of-flight techniques,'' ACI Mater. J. **95**, 168–177 (1998).
[3] Z. P. Bažant and J. Planas, *Fracture and Size Effect in Concrete and Other Quasibrittle Materials* (CRC Press, Boca Raton, 1998).
[4] G. Hevin, O. Abraham, H. Pedersen, and M. Campillo, ''Characterisation of surface cracks with Rayleigh waves: A numerical model,'' NDT & E Int. **31**, 289–297 (1998).
[5] J. S. Popovics, W. J. Song, M. Ghandehari, K. V. Subramaniam, J. D. Achenbach, and S. P. Shah, ''Application of surface wave transmission measurements for crack depth determination in concrete,'' ACI Mater. J. **97**, 127–135 (2000).
[6] W.-J. Song, J. S. Popovics, J. C. Aldrin, and S. P. Shah, ''Measurement of surface wave transmission coefficient across surface-breaking cracks and notches in concrete,'' J. Acoust. Soc. Am. **113**, 717–725 (2003).
[7] J. O. Owino and L. J. Jacobs, ''Attenuation measurements in cement-based materials using laser ultrasonics,'' J. Eng. Mech. **125**, 637–647 (1999).
[8] Y.-C. Jung, W.-B. Na, T. Kundu, and M. Ehsani, ''Damage detection in concrete using Lamb waves,'' *Proceedings of SPIE—The International Society for Optical Engineering* **3995**, pp. 448–458, 2000.
[9] P. Anugonda, J. S. Wiehn, and J. A. Turner, ''Diffusion of ultrasound in concrete,'' Ultrasonics **39**, 429–435 (2001).
[10] Y. Kane, Master's thesis, University of Nebraska-Lincoln, 2003.

On Rayleigh waves in incompressible orthotropic elastic solids

Ray W. Ogden[a)]
Department of Mathematics, University of Glasgow, Glasgow G12 8QW, United Kingdom

Pham Chi Vinh
Faculty of Mathematics, Mechanics and Informatics, Hanoi National University, 334 Nguyen Trai Street, Thanh Xuan, Hanoi, Vietnam

(Received 7 June 2003; accepted for publication 4 November 2003)

In this paper the secular equation for the Rayleigh wave speed in an incompressible orthotropic elastic solid is obtained in a form that does not admit spurious solutions. It is then shown that inequalities on the material constants that ensure positive definiteness of the strain–energy function guarantee existence and uniqueness of the Rayleigh wave speed. Finally, an explicit formula for the Rayleigh wave speed is obtained. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1636464]

PACS numbers: 43.20.Jr, 43.20.Bi, 43.20.Gp [JGH] Pages: 530–533

I. INTRODUCTION

Rayleigh (surface) waves were first studied by Rayleigh (1885) for a compressible isotropic elastic solid. The extension of surface wave analysis to anisotropic elastic materials has been the subject of many studies; see, for example, Stoneley (1963); Chadwick and Smith (1977); Royer and Dieulesaint (1984); Mozhaev (1995); Destrade (2001a); Ting (2002a, c), and references contained therein. Some recent work has focused on incompressible anisotropic elastic solids (Nair and Sotiropoulos, 1997, 1999; Destrade, 2001b; Destrade *et al.*, 2002), while application to materials, compressible or incompressible, subject to prestress has also attracted considerable attention (Dowaikh and Ogden, 1990, 1991; Chadwick, 1997, for example).

Since Green's functions for many elastodynamic problems for a half-space require the solution of the secular equation for Rayleigh waves, formulas for the Rayleigh wave speed in various elastic media are clearly of practical as well as theoretical interest. A formula for the Rayleigh wave speed in compressible isotropic solids was first obtained by Rahman and Barber (1995) for a limited range of values of the parameter $\epsilon \equiv \mu/(\lambda+2\mu)$, where λ and μ are the Lamé constants. For any range of values of ϵ a formula was obtained by Nkemzi (1997); see also Malischewsky (2000). Recently, for some special cases of compressible monoclinic materials with symmetry plane $x_3=0$, formulas for the Rayleigh wave speed were found by Ting (2002b) and Destrade (2003) as the roots of quadratic equations.

Rayleigh waves in incompressible orthotropic elastic materials were examined recently by Destrade (2001b). Destrade used the method of first integrals proposed by Mozhaev (1995) and found a form of the secular equation. He used this to prove that Rayleigh waves exist and are unique in these materials for all values of the relevant material constants. However, the form of the secular equation obtained by use of Mozhaev's method necessarily admits spurious solutions. Thus, the analysis of Destrade requires some modification. The secular equation for Rayleigh waves in incompressible orthotropic materials presented recently by Destrade *et al.* (2002) also admits spurious solutions.

The aim of the present paper is to obtain a formula for the Rayleigh wave speed in an incompressible orthotropic elastic material. For this purpose a form of the secular equation that does not admit spurious solutions is required.

In Sec. II the basic equations and notation are presented for describing motion in an incompressible orthotropic elastic material. We consider a half-space whose boundary is a symmetry plane of the material. Since the equations for time-harmonic waves propagating parallel to the boundary of this half-space decouple into a plane motion, in the plane defined by the half-space normal and the direction of propagation, and a motion normal to that plane (see, for example, Destrade, 2001b), it suffices to consider the plane strain case. In Sec. III the secular equation for Rayleigh waves is derived in the desired form and it is shown how it relates to the form given by Destrade (2001b). Existence and uniqueness results are presented in Sec. IV. In particular, it is shown that Rayleigh waves exist and are unique in an incompressible orthotropic elastic solid, provided the inequalities

$$\gamma \equiv c_{66} > 0, \quad \delta \equiv c_{11} + c_{22} - 2c_{12} > 0 \tag{1}$$

are satisfied, where c_{11}, c_{12}, c_{22}, and c_{66} are material constants associated with the considered plane, which, in Cartesian coordinates, is taken to be the (x_1,x_2) plane. These inequalities are necessary and sufficient for the strain energy (specialized to the considered plane) to be positive definite. It is also easy to show that they are necessary and sufficient for strong ellipticity to hold for the considered motion.

An explicit formula for the Rayleigh speed is derived in Sec. V by using the theory of cubic equations. Corresponding results for the compressible theory will be discussed in a separate paper.

II. BASIC EQUATIONS

Let (x_1,x_2,x_3) be Cartesian coordinates and consider an orthotropic elastic material occupying the half-space $x_2<0$, with traction-free boundary $x_2=0$, which is a plane of sym-

a)Electronic mail: rwo@maths.gla.ac.uk

metry for the orthotropy. We consider a plane motion in the (x_1,x_2) plane with displacement components (u_1,u_2,u_3) such that

$$u_i=u_i(x_1,x_2,t), \quad i=1,2, \quad u_3\equiv 0, \tag{2}$$

where t is time. The nonzero components ϵ_{ij} of the infinitesimal strain tensor are given by

$$\epsilon_{ij}=\tfrac{1}{2}(u_{i,j}+u_{j,i}), \quad i,j=1,2, \tag{3}$$

where a comma signifies partial differentiation with respect to spatial variables.

For an incompressible material, we have

$$u_{1,1}+u_{2,2}=0, \tag{4}$$

from which we deduce the existence of a scalar function, denoted $\psi(x_1,x_2,t)$, such that

$$u_1=\psi_{,2}, \quad u_2=-\psi_{,1}. \tag{5}$$

For the considered motion the relevant components of the stress are given by, for example, Nair and Sotiropoulos (1997). Thus,

$$\sigma_{11}=-p+c_{11}\epsilon_{11}+c_{12}\epsilon_{22},$$
$$\sigma_{22}=-p+c_{12}\epsilon_{11}+c_{22}\epsilon_{22}, \quad \sigma_{12}=2c_{66}\epsilon_{12}, \tag{6}$$

where σ_{ij}, $i=1, 2$, are components of the stress tensor, c_{ij} are elastic constants of the material in standard notation, and $p=p(x_1,x_2,t)$ is the hydrostatic pressure associated with the incompressibility constraint. Note that, in general, $\sigma_{33}\neq 0$, but we shall not need to use this component of stress.

For the strain energy to be positive definite for the considered plane motion, the elastic constants c_{ij} in Eq. (6) must satisfy the inequalities given in Eq. (1).

In the absence of body forces the relevant components of the equation of motion are

$$\sigma_{11,1}+\sigma_{12,2}=\rho\ddot{u}_1, \quad \sigma_{12,1}+\sigma_{22,2}=\rho\ddot{u}_2, \tag{7}$$

where dots denote partial differentiation with respect to t, and ρ is the mass density of the material.

Use of Eqs. (3), (5), and (6) in Eq. (7) and elimination of p by cross differentiation leads to an equation for ψ, namely

$$\gamma\psi_{,1111}+2\beta\psi_{,1122}+\gamma\psi_{,2222}=\rho(\ddot{\psi}_{,11}+\ddot{\psi}_{,22}), \tag{8}$$

where

$$2\beta=\delta-2\gamma, \tag{9}$$

and γ and δ are defined by Eq. (1).

In terms of the stress components the traction-free boundary conditions are written

$$\sigma_{12}=\sigma_{22}=0 \quad \text{on } x_2=0. \tag{10}$$

Using Eqs. (3), (5), (6), and the first of (7), Eq. (10) can be expressed as conditions on ψ. This requires differentiation of σ_{22} with respect to x_1 so as to facilitate elimination of the term in p, as in Dowaikh and Ogden (1990) for Rayleigh waves on a prestrained half-space of incompressible isotropic elastic material. The resulting boundary conditions are

$$\gamma(\psi_{,22}-\psi_{,11})=0,$$
$$\gamma(\psi_{,222}-\psi_{,112})+\delta\psi_{,112}-\rho\ddot{\psi}_{,2}=0 \quad \text{on } x_2=0. \tag{11}$$

We shall also require that

$$\psi(x_1,x_2,t)\to 0 \quad \text{as } x_2\to-\infty. \tag{12}$$

III. RAYLEIGH WAVES: SECULAR EQUATION

We now consider harmonic waves propagating in the x_1 direction, and we write ψ in the form

$$\psi(x_1,x_2,t)=\phi(y)\exp[ik(x_1-ct)], \tag{13}$$

where k is the wave number, c is the wave speed, $y=kx_2$, and the function ϕ is to be determined.

Substitution of Eq. (13) into Eq. (8) yields

$$\gamma\phi''''-(2\beta-\rho c^2)\phi''+(\gamma-\rho c^2)\phi=0, \tag{14}$$

where, in Eq. (14) and the following, a prime on ϕ indicates differentiation with respect to y.

In terms of ϕ the boundary conditions (11) become

$$\phi''(0)+\phi(0)=0, \quad \gamma\phi'''(0)+(\gamma-\delta+\rho c^2)\phi'(0)=0, \tag{15}$$

in the first of which we have omitted the factor γ on the assumption that $\gamma\neq 0$. From Eq. (12) we also require that

$$\phi(x_2)\to 0 \quad \text{as } x_2\to-\infty. \tag{16}$$

Thus, the problem is reduced to solving Eq. (14) with the boundary conditions (15) and (16). The general solution for $\phi(y)$ that satisfies the condition (16) is

$$\phi(y)=A\exp(s_1y)+B\exp(s_2y), \tag{17}$$

where A and B are constants, while s_1 and s_2 are the solutions of the equation

$$\gamma s^4-(2\beta-\rho c^2)s^2+(\gamma-\rho c^2)=0, \tag{18}$$

with positive real parts.

From Eq. (18) it follows that

$$s_1^2+s_2^2=(2\beta-\rho c^2)/\gamma, \quad s_1^2s_2^2=(\gamma-\rho c^2)/\gamma. \tag{19}$$

If the roots s_1^2 and s_2^2 of the quadratic Eq. (18) are real, then they must be positive to ensure that s_1 and s_2 can have a positive real part. If they are complex then they are conjugate. In either case the product $s_1^2s_2^2$ must be positive and hence a real (surface) wave speed c satisfies the inequalities

$$0<\rho c^2<\gamma. \tag{20}$$

Note that the limiting wave speed such that $\rho c^2=\gamma$ is the speed of a shear body wave, not a surface wave.

Substituting Eq. (17) into the boundary conditions (15), we obtain the equations

$$(1+s_1^2)A+(1+s_2^2)B=0,$$
$$[\gamma(s_1^2+1)+\rho c^2-\delta]s_1A+[\gamma(s_2^2+1)+\rho c^2-\delta]s_2B=0, \tag{21}$$

for A and B. For nontrivial solution the determinant of coefficients of the system (21) must vanish. After removal of the factor (s_1-s_2), this yields

$$\gamma(s_1^2+s_2^2+s_1^2s_2^2)+(\delta-\rho c^2)s_1s_2+\gamma-\delta+\rho c^2=0. \tag{22}$$

Use of Eq. (19) in Eq. (22) then leads to

$$(\delta-\rho c^2)\sqrt{1-\rho c^2/\gamma}-\rho c^2=0, \tag{23}$$

which is the required secular equation for the wave speed through ρc^2. Note that for this equation to have a real solution for c it is necessary, in addition to Eq. (20), that the inequality

$$\rho c^2 < \delta, \tag{24}$$

must hold. Note that a solution of Eq. (22) with $s_1 = s_2$ is admissible, but it can be shown that this corresponds to a real surface wave if and only if $\delta = 32\gamma/9$ and $\rho c^2 = \delta/4$, a solution that is given by (23) for this very restricted combination of material constants. Of course, in this case, the solution (17) requires modification.

By rearranging Eq. (23) and squaring to eliminate the square root, we obtain the secular equation derived by Destrade (2001b), which, in the present notation, can be written

$$(\delta - \rho c^2)^2 (1 - \rho c^2/\gamma) = (\rho c^2)^2. \tag{25}$$

Destrade (2001b) used the method of first integrals proposed by Mozhaev (1995) and needed to impose the restriction $\rho c^2 \neq \delta$, which in our derivation is automatically satisfied through Eq. (24). Using this equation, Destrade concluded that a unique Rayleigh wave exists in an incompressible orthotropic elastic material for any values of γ and δ. However, Eq. (25) may have spurious solutions for ρc^2 that are not solutions of Eq. (23), and it is therefore advisable to avoid drawing conclusions on the basis of Eq. (25).

IV. EXISTENCE AND UNIQUENESS OF RAYLEIGH WAVES

We now show that the inequalities $\gamma > 0$ and $\delta > 0$ jointly ensure the existence and uniqueness of a Rayleigh wave. For this purpose it is convenient to introduce the new variable $\eta = \sqrt{1 - \rho c^2/\gamma}$ so that the secular equation (23) may be rewritten as

$$f(\eta) \equiv \eta^3 + \eta^2 + (\delta/\gamma - 1)\eta - 1 = 0, \quad 0 < \eta < 1. \tag{26}$$

Then

$$f(0) = -1 < 0, \quad f(1) = \delta/\gamma > 0, \tag{27}$$

which guarantees that Eq. (26) has at least one solution in the interval (0,1).

We also have

$$f'(\eta) = 3\eta^2 + 2\eta + \delta/\gamma - 1, \quad f''(\eta) > 0 \quad (\eta > 0). \tag{28}$$

If $\delta \geqslant \gamma$ then it follows that $f'(\eta) > 0$ for $\eta > 0$ and hence f is monotonic increasing for $\eta > 0$. In this case the solution for η is unique. If, on the other hand, $0 < \delta < \gamma$ then $f'(0) < 0$. Thus, f has a maximum for $\eta < 0$ and a minimum for $\eta > 0$. By the inequality in Eq. (28) f therefore decreases to a minimum as η increases from 0, and thereafter increases monotonically. Hence, the solution is also unique in this case.

We therefore conclude that in an incompressible orthotropic elastic half-space there exists a unique Rayleigh wave provided the material constants satisfy the conditions (1), which ensure that the strain–energy function is positive definite for the considered plane strain restriction. We note in passing that if $\delta \leqslant 0$ then it can be seen immediately that Eq. (23) has no real nonzero solution for c whatever the sign of γ, although it is not physically meaningful to admit nonpositive values of these constants.

V. A FORMULA FOR THE WAVE SPEED

In this section we derive an explicit formula for the wave speed, given that $\gamma > 0$, $\delta > 0$, by seeking the unique root, η_0 say, of Eq. (26) in the interval (0,1). The wave speed c is then given by

$$\rho c^2 = \gamma(1 - \eta_0^2). \tag{29}$$

We now show that the cubic Eq. (26) has only one real root, namely η_0, the other two being complex.

According to the theory of cubic equation (see, for example, Cowles and Thompson, 1947 or Abramowitz and Stegun 1974), the nature of the three roots of the cubic

$$\eta^3 + a_2\eta^2 + a_1\eta + a_0 = 0, \tag{30}$$

is determined by the sign of the discriminant D defined by

$$D = R^2 + Q^3, \tag{31}$$

where R and Q are given in terms of the coefficients a_0, a_1, a_2 by

$$R = \tfrac{1}{54}(9a_1a_2 - 27a_0 - 2a_2^3), \quad Q = \tfrac{1}{9}(3a_1 - a_2^2). \tag{32}$$

If $D > 0$, Eq. (30) has one real root and two complex conjugate roots. If $D = 0$, the equation has three real roots, at least two of which are equal. If $D < 0$, Eq. (30) has three distinct real roots. In the first case ($D > 0$) the single real root η_0 is given by Cardano's formula (Cowles and Thompson, 1947; Abramowitz and Stegun, 1974) in the form

$$\eta_0 = -\tfrac{1}{3}a_2 + (R + \sqrt{D})^{1/3} + (R - \sqrt{D})^{1/3}. \tag{33}$$

For the secular equation in the form Eq. (26), we have

$$a_0 = -1, \quad a_1 = \Delta - 1, \quad a_2 = 1, \tag{34}$$

and hence

$$R = \frac{\Delta}{6} + \frac{8}{27}, \quad Q = \frac{\Delta}{3} - \frac{4}{9}, \tag{35}$$

where $\Delta = \delta/\gamma$. Using Eq. (35) in Eq. (31), it is easy to verify that

$$D = \tfrac{1}{108}\Delta(4\Delta^2 - 13\Delta + 32). \tag{36}$$

It is clear from Eq. (36) that $D > 0$ provided $\Delta > 0$. Thus, Eq. (30) has only one real root, necessarily within the required range of values.

Use of Eqs. (34), (35), and (36) in Eq. (33) leads to

$$\eta_0 = \tfrac{1}{3}\Big[-1 + \sqrt[3]{[9\Delta + 16 + 3\sqrt{3}\sqrt{\Delta(4\Delta^2 - 13\Delta + 32)}]/2} + \sqrt[3]{[9\Delta + 16 - 3\sqrt{3}\sqrt{\Delta(4\Delta^2 - 13\Delta + 32)}]/2}\Big]. \tag{37}$$

From Eqs. (29) and (37) the speed c of the Rayleigh wave is given by

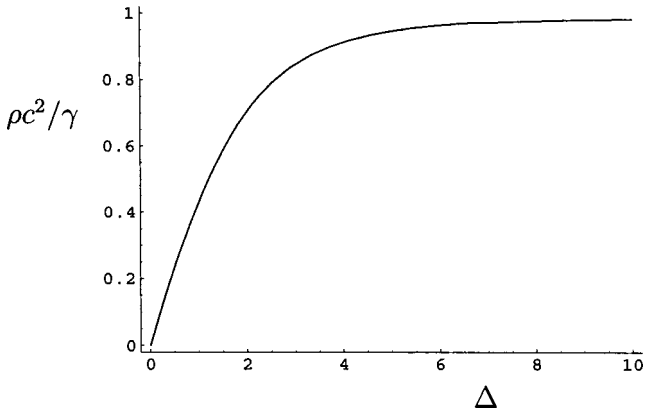


FIG. 1. Plot of $\rho c^2/\gamma$ against $\Delta(>0)$ from Eq. (38).

$$\rho c^2/\gamma = 1 - \tfrac{1}{9}[-1 + \sqrt[3]{[9\Delta + 16 + 3\sqrt{3}\sqrt{\Delta(4\Delta^2 - 13\Delta + 32)}]/2} + \sqrt[3]{[9\Delta + 16 - 3\sqrt{3}\sqrt{\Delta(4\Delta^2 - 13\Delta + 32)}]/2}]^2. \tag{38}$$

For an (incompressible) isotropic material $c_{11}=c_{22}$, $c_{11}-c_{12}=2\mu$, and $c_{66}=\mu$, where μ is the classical shear modulus, and hence, by Eq. (1), $\Delta=4$. In this case the formula (38) specializes to

$$\rho c^2/\gamma = 1 - \tfrac{1}{9}[\sqrt[3]{6\sqrt{33}+26} - \sqrt[3]{6\sqrt{33}-26} - 1]^2. \tag{39}$$

This is approximately 0.9126, which is the classical value for an incompressible isotropic elastic solid (see, for example, Ewing *et al.*, 1957).

In Fig. 1 a plot of $\rho c^2/\gamma$ against $\Delta(>0)$ based on Eq. (38) is shown in order to illustrate the dependence of the wave speed on the ratio of material constants. The wave speed is very small for small Δ and increases rapidly as Δ increases, reaching its isotropic value for $\Delta=4$ and then approaching an asymptotic value with $\rho c^2/\gamma \to 1$ as Δ becomes very large. The asymptotic limit corresponds to a wave speed equal to the shear wave speed. Note that δ may be interpreted as a shear modulus of the material; indeed, in the isotropic case $\delta=2\mu$, where μ is the Lamé shear modulus. Thus, the limit $\Delta \to 0$ (which is not applicable for isotropic materials) corresponds to a material with one vanishingly small shear modulus. Similarly, γ is a shear modulus and, if $\delta \neq 0$, in the limit $\gamma \to 0$ we have $\Delta \to \infty$. Thus, we have interpretations for the two extreme values of Δ.

ACKNOWLEDGMENTS

The work is partly supported by the Ministry of Education and Training of Vietnam and completed during a visit of the second author to the Department of Mathematics, University of Glasgow, UK.

Abramowitz, M., and Stegun, I. A. (**1974**). *Handbook of Mathematical Functions* (Dover, New York).

Chadwick, P. (**1997**). "The application of the Stroh formulation to prestressed elastic media," Math. Mech. Solids **2**, 379–403.

Chadwick, P., and Smith, G. D. (**1977**). "Foundations of the theory of surface waves in anisotropic elastic materials," Adv. Appl. Mech. **17**, 303–376.

Cowles, W. H., and Thompson, J. E. (**1947**). *Algebra* (Van Nostrand, New York).

Destrade, M. (**2001a**). "The explicit secular equation for surface acoustic waves in monoclinic elastic crystals," J. Acoust. Soc. Am. **109**, 1398–1402.

Destrade, M. (**2001b**). "Surface waves in orthotropic incompressible materials," J. Acoust. Soc. Am. **110**, 837–840.

Destrade, M. (**2003**). "Rayleigh waves in symmetry planes of crystals: explicit secular equations and some explicit wave speeds," Mech. Mater. **35**, 931–939.

Destrade, M., Martin, P. A., and Ting, T. C. T. (**2002**). "The incompressible limit in linear anisotropic elasticity, with application to surface waves and elastostatics," J. Mech. Phys. Solids **50**, 1453–1468.

Dowaikh, M. A., and Ogden, R. W. (**1990**). "On surface waves and deformations in a pre-stressed incompressible elastic solid," IMA J. Appl. Math. **44**, 261–284.

Dowaikh, M. A., and Ogden, R. W. (**1991**). "On surface waves and deformations in a compressible elastic half-space," Stab. Appl. Anal. Cont. Media **1**, 27–45.

Ewing, W. M., Jardetzky, W. F., and Press, F. (**1957**). *Elastic Waves in Layered Media* (McGraw-Hill, New York).

Malischewsky, P. G. (**2000**). Comment on "A new formula for the velocity of Rayleigh waves by D. Nkemzi," Wave Motion **31**, 93–96.

Mozhaev, V. G. (**1995**). "Some new ideas in the theory of surface acoustic waves in anisotropic media," in *IUTAM Symposium on Anisotropy, Inhomogeneity and Nonlinearity in Solid Mechanics*, edited by D. F. Parker and A. H. England (Kluwer, Dordrecht), pp. 455–462.

Nair, S., and Sotiropoulos, D. A. (**1997**). "Elastic waves in orthotropic incompressible materials and reflection from an interface," J. Acoust. Soc. Am. **102**, 102–109.

Nair, S., and Sotiropoulos, D. A. (**1999**). "Interfacial waves in incompressible monoclinic materials with an interlayer," Mech. Mater. **31**, 225–233.

Nkemzi, D. (**1997**). "A new formula for the velocity of Rayleigh waves," Wave Motion **26**, 199–205.

Rahman, M., and Barber, J. R. (**1995**). "Exact expressions for the roots of the secular equation for Rayleigh waves," J. Appl. Mech. **62**, 250–252.

Rayleigh, Lord (**1885**). "On waves propagated along the plane surface of an elastic solid," Proc. R. Soc. London, Ser. A **17**, 4–11.

Royer, D., and Dieulesaint, E. (**1984**). "Rayleigh wave velocity and displacement in orthorhombic, tetragonal, hexagonal and cubic crystals," J. Acoust. Soc. Am. **75**, 1438–1444.

Stoneley, R. (**1963**). "The propagation of surface waves in an elastic medium with orthorhombic symmetry," Geophys. J. R. Astron. Soc. **8**, 176–186.

Ting, T. C. T. (**2002a**). "An explicit secular equation for surface waves in an elastic material of general anisotropy," Q. J. Mech. Appl. Math. **55**, 297–311.

Ting, T. C. T. (**2002b**). "A unified formalism for elastostatics or steady state motion of compressible or incompressible anisotropic elastic materials," Int. J. Solids Struct. **39**, 5427–5445.

Ting, T. C. T. (**2002c**). "Explicit secular equations for surface waves in monoclinic materials with symmetry plane at $x_1=0$, $x_2=0$ or $x_3=0$," Proc. R. Soc. London, Ser. A **458**, 1017–1031.

On the calculation of the transmission line parameters for long tubes using the method of multiple scales

Stefan Scheichl[a)]
Acoustics Research Institute, Austrian Academy of Sciences, A-1010 Vienna, Austria

(Received 22 March 2003; revised 10 October 2003; accepted 17 November 2003)

The present paper deals with the classical problem of linear sound propagation in tubes with isothermal walls. The perturbation technique of the method of multiple scales in combination with matched asymptotic expansions is applied to derive the first-order solutions and, in addition, the second-order solutions representing the correction due to boundary layer attenuation. The propagation length is assumed to be so large that in order to obtain asymptotic solutions which extend over the whole spatial range the first-order corrections to the classical attenuation rates of the different modes come into play as well. Starting with the case of the characteristic wavelength being large compared to the characteristic dimension of the duct, the analysis is then extended to the case where both of these quantities are of the same order of magnitude. Furthermore, the transmission line parameters and the transfer functions relating the sound pressures at the ends of the duct to the axial velocities are calculated. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639323]

PACS numbers: 43.20.Mv, 43.20.Bi, 43.20.Hq [MO] Pages: 534–555

I. INTRODUCTION

The subject of linear sound propagation in rigid tubes with isothermal walls has been attracting considerable interest over the years. The ''exact'' solution for tubes of circular cross-sectional shapes derived by Kirchhoff[1] (see also Rayleigh,[2] pp. 319–328) accounts for the effects of shear viscosity and heat conduction on the attenuation of sound waves. Later Zwikker and Kosten[3] (pp. 25–40) and independently also Iberall,[4] Daniels,[5] and Kraak[6] introduced an approximate theory based on the so-called *low reduced frequency assumptions* that enabled the simplification of the basic equations such that the transmission line parameters could be given in closed form. A thorough discussion of the applicability of this approach including a comparison with numerical solutions of Kirchhoff's general dispersion equations is presented in the 1975 paper by Tijdeman.[7] More recently, Stinson[8] considered an alternative treatment of the problem applying simplifying approximations to the equations that make up the Kirchhoff solution, rather than reducing the governing equations, and showed the equivalence of both approaches. These investigations then provided the basis for developing a general procedure applicable to tubes of arbitrary cross-sectional shape. Similar calculations were also carried out by Kergormard.[9]

From the point of view of a perturbation analysis the low reduced frequency assumptions can be interpreted as defining two scaling parameters that relate the most relevant geometrical scales, i.e., the wavelength, the characteristic diameter of the duct, and the thickness of the acoustic boundary layer, to each other: The spatial range consumed by the boundary layer as well as the diameter are presupposed to be small compared to the wavelength. Further simplifications are then possible assuming the boundary layer to be either small or large relative to the tube diameter, which introduces a third scaling parameter and thus two different ordering relationships for the two other parameters. In the following, these cases will be called the *high* and the *low frequency limit*, respectively. If, however, such additional constraints are imposed *a posteriori* on the solutions given by Zwikker and Kosten in order to derive approximate series expansions of the transmission line parameters with respect to that third ratio (see, e.g., Keefe's results[10] for the cylindrical tube), it will remain unclear whether the resulting expressions are the correct asymptotic solutions one would obtain if the two different ordering relationships were applied to the basic equations themselves.

Moreover, the length of the duct might become so large that the exponentially growing effects arising from viscosity and heat conduction in the boundary layer do not only affect the second-order terms of the sound pressure but also the leading order terms. The present study is motivated by the observation that sound propagation in tubes of this type has not yet been systematically studied. Thus, one of its primary aims is to derive the asymptotically correct solutions for the involved field quantities including the second-order terms that extend over a considerably large spatial range, suggesting the application of the *method of multiple scales*, as presented, e.g., in Nayfeh[11] or Crighton *et al.*[12] (pp. 209–232). By the removal of secular terms, the extra freedom such an approach introduces can be exploited to increase the range of validity of the asymptotic expansions. Since the following calculations proceed from the assumption that the boundary layer is small compared to the tube diameter (high frequency limit), the changes in lateral direction will be analyzed using the perturbation technique of the *method of matched asymptotic expansions*. A further goal to be pursued is the derivation of the asymptotically correct expressions for the transmission line parameters and the transfer functions of a long tube up to the second-order terms. The investigations are structured as follows.

As far as the diameter to wavelength ratio is concerned,

a)Electronic mail: stefan.scheichl@oeaw.ac.at

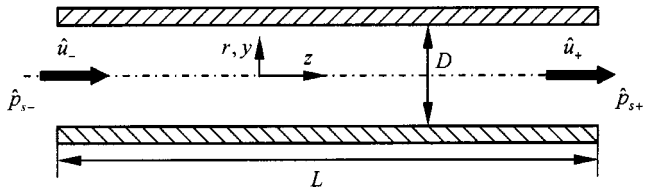


FIG. 1. Sketch and notation of an acoustical four pole.

the first part of the analysis deals with long wavelengths, in accordance with the low reduced frequency assumptions. Results for the first-order and the second-order terms of the sound pressure, the velocity components, and the transfer matrix in the case of long circular tubes will be presented. Similar solutions derived for the case of long rectangular and slit-shaped tubes are found in the appendices; these are, however, valid only to leading order. In the second part, beginning with Sec. V, the long wavelength assumption is then replaced with the condition that the characteristic wavelength and the diameter of the tube are of the same order of magnitude. To demonstrate the utility of the procedure, the transfer functions up to the second order of a long circular tube will again be derived. Since in this case the occurrence of higher order modes has to be taken into account as well, the study will confine itself to the case of axisymmetric flow. It should be mentioned that in a similar investigation concerning the sound propagation in a slit-shaped waveguide carried out by Anderson and Vaidya[13] the authors pointed out that the application of the method of multiple scales to the linear problem requires several observations suggested by results obtained from the so-called classical analysis that poses the boundary value problem as an eigenvalue problem, and would fail otherwise. However, as it will turn out in the following investigation, here such difficulties are not encountered.

II. PROBLEM FORMULATION

A. Transmission line parameters for acoustical four poles

For the time being, assume that the driving frequency is sufficiently low that only the fundamental mode is able to propagate in a tube of length L regarded as a transmission line (see Fig. 1). Furthermore, let $\hat{Z}$ and $\hat{Y}$ be the series impedance and shunt admittance per unit length along the z axis. The sound pressure $p_s(z,t)=\Re(\hat{p}_s(z)e^{j\omega t})$ and the volume flow $u(z,t)=\Re(\hat{u}(z)e^{j\omega t})$ are then given by

$$\frac{d\hat{p}_s}{dz}=-\hat{Z}\hat{u}, \quad \frac{d\hat{u}}{dz}=-\hat{Y}\hat{p}_s. \tag{1}$$

Hence, the four-pole transfer matrix $\hat{\mathbf{A}}$ can be formulated as

$$\hat{\mathbf{A}}=\begin{bmatrix} \cosh(\hat{\Gamma}L) & \hat{Z}_c\sinh(\hat{\Gamma}L) \\ \dfrac{1}{\hat{Z}_c}\sinh(\hat{\Gamma}L) & \cosh(\hat{\Gamma}L) \end{bmatrix}, \tag{2}$$

where

$$\hat{Z}_c=\sqrt{\frac{\hat{Z}}{\hat{Y}}}, \quad \hat{\Gamma}=\sqrt{\hat{Z}\hat{Y}} \tag{3}$$

are the characteristic impedance and the propagation parameter of the duct. The values of the sound pressure and the volume flow at the entrance of the tube, in the following denoted by $\hat{p}_{s-}$ and $\hat{u}_-$, respectively, can then simply be calculated from the values $\hat{p}_{s+}$ and $\hat{u}_+$ at the end of the tube by employing the relationship

$$\begin{pmatrix} \hat{p}_{s-} \\ \hat{u}_- \end{pmatrix}=\hat{\mathbf{A}}\begin{pmatrix} \hat{p}_{s+} \\ \hat{u}_+ \end{pmatrix}. \tag{4}$$

B. Low reduced frequency and low Mach number assumptions

In order to be able to calculate the transmission line parameters entering the transfer matrix from the basic equations, i.e., the two- or three-dimensional Navier–Stokes equations, the energy equation, the continuity equation, and the equation of state for a perfect gas, the following so-called low reduced frequency assumptions (see, e.g., Tijdeman[7]) are adopted:

$$\mathrm{Re}=\frac{c_0\lambda\rho_0}{\mu_0}\gg 1, \quad l=\frac{D}{\lambda}\ll 1. \tag{5}$$

Here, the quantities c_0, λ, ρ_0, μ_0, D, and the parameters Re and l denote the speed of sound, the characteristic wavelength, the density of the fluid, the dynamic shear viscosity, the characteristic dimension of the cross section, the acoustic Reynolds number, and the reduced frequency, which is of the order of the Helmholtz number $\mathrm{He}=\omega D/(2c_0)$. By the subscript 0, quantities evaluated at the equilibrium reference state are indicated.

Since the Prandtl number

$$\mathrm{Pr}=\frac{\mu_0 C_p}{\kappa_0}=O(1), \tag{6}$$

where C_p and κ_0 represent the specific heat at constant pressure and the thermal conductivity, is of order $O(1)$ for a wide class of fluid (e.g., Pr≈0.7 in case of air), the thermal as well as the viscous boundary layer thicknesses are given by (see, e.g., Morse and Ingard,[14] p. 286)

$$\delta\sim\sqrt{\frac{\mu_0\lambda}{c_0\rho_0}}=\frac{\lambda}{\sqrt{\mathrm{Re}}}. \tag{7}$$

Consequently, from the first restriction in Eq. (5) it follows that δ is small compared to the characteristic wavelength, ensuring that the flow is not dominated by viscous effects; it is easily verified that $\mathrm{Re}\gg 1$ holds for the complete range of audible and even a wide range of ultrasonic frequencies, provided that the fluid is air. This, together with the long wavelength assumption $l\ll 1$ stated in Eq. (5), then guarantees that only a single mode propagates over large distances relative to the tube diameter and a simplified, one-dimensional formulation of the problem as in Eq. (1) can be derived.

In the following, it will furthermore be assumed that the Mach number

$$M=\frac{W}{c_0}\ll 1, \tag{8}$$

with W as a characteristic particle velocity in longitudinal direction, is small as well, so that a linear analysis of the problem is possible.

C. Ducts with circular cross sections and isothermal walls

By imposing the low reduced frequency assumptions ($D=2R$), evaluation of the linearized basic equations yields (see Zwikker and Kosten,[3] pp. 25–40, but also Refs. 4–7, 10)

$$\hat{Z}=\frac{j\omega\rho_0}{S\left(1-2\sqrt{\frac{\mu_0}{-j\omega\rho_0}}\frac{\hat{G}\left(R\sqrt{\frac{-j\omega\rho_0}{\mu_0}}\right)}{R}\right)},$$
$$\hat{Y}=\frac{j\omega S\left[1+2(\gamma-1)\sqrt{\frac{\kappa_0}{-j\omega\rho_0 C_p}}\frac{\hat{G}\left(R\sqrt{\frac{-j\omega\rho_0 C_p}{\kappa_0}}\right)}{R}\right]}{\gamma p_0}, \tag{9}$$

where

$$\hat{G}(\hat{\xi})=\frac{\mathrm{J}_1(\hat{\xi})}{\mathrm{J}_0(\hat{\xi})} \tag{10}$$

and the cross-sectional area $S=R^2\pi$. The quantities p_0 and $\gamma=C_p/C_v$ are the equilibrium pressure and the ratio of the specific heats. Worth mentioning is the fact that Eq. (9) can also be derived by averaging the expressions for the velocity in the direction of the tube axis and the sound pressure given in Morse and Ingard[14] (pp. 519–522) over the cross section. These solutions were obtained from an axisymmetric two-dimensional analysis by using assumptions very similar to those Stinson's generalized theory for tubes of arbitrary cross-sectional shape is based on (see Ref. 8 and Appendix A).

High and low frequency limits can now be defined as

$$\text{LFL:}\quad l^2\,\mathrm{Re}\sim\left(\frac{R}{\delta}\right)^2\sim\frac{\omega\rho_0R^2}{\mu_0}=\mathrm{St}^2\ll 1,$$
$$\text{HFL:}\quad \mathrm{St}^2\gg 1. \tag{11}$$

Here, the quantity St denotes the so-called Stokes number, its inverse is sometimes referred to as shear wave number Sh. As mentioned earlier, $\mathrm{Pr}=O(1)$ and therefore the LFL and HFL can easily be deduced from Eq. (9) by applying a power series expansion with respect to St and St^{-1}, respectively. In connection with the HFL, it should be noted that in the limit as $\mathrm{St}\to\infty$,

$$\hat{G}(\hat{\xi})\to -j, \tag{12}$$

as $\hat{\xi}$ is proportional to $\mathrm{St}\sqrt{-j}$.

Consequently, the expressions (9) reduce to

$$\text{LFL:}\quad \hat{Z}=\frac{8\mu_0}{R^2S}\left(1+j\frac{\omega\rho_0R^2}{6\mu_0}+O(\mathrm{St}^4)\right),$$
$$\hat{Y}=\frac{\omega S}{p_0}\left(j+\frac{\gamma-1}{\gamma}\frac{\omega\rho_0R^2}{8\mu_0}\frac{\mu_0C_p}{\kappa_0}+O(\mathrm{St}^4)\right), \tag{13}$$

which corresponds to Rayleigh's *narrow* tube solution (Ref. 2, p. 327), if terms of $O(\mathrm{St}^2)$ are neglected as well, and

$$\text{HFL:}\quad \hat{Z}=\frac{\omega\rho_0}{S}\left[j+(1+j)\sqrt{\frac{2\mu_0}{\omega\rho_0R^2}}+O(\mathrm{St}^{-2})\right],$$
$$\hat{Y}=\frac{\omega S}{\gamma p_0}\left[j+(1+j)(\gamma-1)\sqrt{\frac{2\mu_0}{\omega\rho_0R^2}\frac{\kappa_0}{\mu_0C_p}}+O(\mathrm{St}^{-2})\right], \tag{14}$$

which is in accordance with Kirchhoff's *wide* tube solution.[1] Simplified expressions for the limiting values of $\hat{Z}$ can also be found in the book by Beranek[15] (pp. 135–138). Equivalent results for the HFL of the series impedance as well as the shunt admittance in case of rectangular or slit-shaped cross sections are given in Appendices A and B.

A completely different method of finding the LFL and HFL would be an asymptotic analysis of the basic equations themselves, with M, Re^{-1}, l, and either St or St^{-1} used as (small) perturbation parameters. In the case of the HFL such an approach then necessitates separate investigations of the acoustic motion in the core region and in the boundary layer, since the scaling of the terms in the basic equations changes completely, depending on which region is under consideration. Such a so-called matched asymptotic analysis, which again reproduces the solutions (14), is presented, e.g., in Makarov and Vatrushina[16] as well as in Qi *et al.*[17] However, if the evolution of waves over distances of the order $O(\mathrm{St}\,\lambda)$ is taken into consideration, evaluation of the transfer matrix $\hat{\mathbf{A}}$ as defined in Eq. (2) using $\hat{Z}$ and $\hat{Y}$ from Eq. (14) will produce results which are valid only to leading order. This is a direct consequence of the fact that the exponential terms of the order $O(e^{(\mathrm{St}^{-1}\lambda^{-1}L)})$ contained in the transfer matrix will then become order $O(1)$ quantities. In other words, in order to calculate asymptotically correct expressions for the leading order terms and the correction terms [of order $O(\mathrm{St}^{-1})$] of the quantities $p_s(z,t)$ and $u(z,t)$, the above-presented results for the HFL have to be based on the assumption that the propagation length L of the acoustic waves is comparable to the wavelength.

The aim of the following investigations is thus twofold: First, to show that the application of the method of multiple scales (MMS) in the HFL together with a matched asymptotic analysis leads to analytical solutions for the sound pressure and the volume flow *including* the second-order terms that are uniformly valid over a considerably larger spatial range than that constituted by the wavelengths and, second, to derive the asymptotically correct expressions for the coefficients of the transfer matrix $\hat{\mathbf{A}}$. As will furthermore be shown in Sec. V, the MMS can even be applied if the condition $l\ll 1$ is relaxed such that the reduced frequency

is assumed to be of order $O(1)$ and, consequently, the excitation of higher order modes can no longer be disregarded.

III. BASIC EQUATIONS

A natural nondimensionalization of the governing equations involves the wavelength λ, the radius R, as well as the equilibrium quantities p_0, ρ_0, c_0, and μ_0 introduced above. Nondimensional variables are then constructed from

$$z^*=\frac{z}{\lambda},\quad L^*=\frac{L}{\lambda},\quad r^*=\frac{r}{R},\quad t^*=\frac{tc_0}{\lambda},\quad \omega^*=\frac{\omega\lambda}{c_0},$$

$$v_z^*=\frac{v_z}{c_0},\quad v_r^*=\frac{v_r}{c_0},\quad u^*=\frac{u}{c_0 S},\quad p^*=\frac{p}{\gamma p_0},\quad \rho^*=\frac{\rho}{\rho_0},\tag{15}$$

$$\vartheta^*=\frac{\vartheta}{\vartheta_0},\quad Z^*=\frac{Zc_0\lambda S}{\gamma p_0},\quad Y^*=\frac{Y\gamma p_0\lambda}{c_0 S},\quad \eta^*=\frac{\eta_0}{\mu_0}.$$

Here v_z, v_r, p, ϑ, and η denote, respectively, the velocities in axial and radial direction, the fluid pressure, the temperature, and the bulk viscosity. Furthermore, the scaling parameters

$$\epsilon=a\frac{R}{\lambda},\quad \alpha=d\frac{\lambda}{R\sqrt{\mathrm{Re}}}\tag{16}$$

are introduced where a and d are arbitrary constants of order $O(1)$, which, together with Eqs. (5) and (11), leads to the relationships $\epsilon\sim l\sim \mathrm{He}$ and $\alpha\sim \mathrm{St}^{-1}$.

In the following analysis it will be assumed that the variations of the thermal conductivity and the dynamic viscosities are so small that these quantities can be regarded as constant, i.e., $\kappa=\kappa_0$, $\mu=\mu_0$, and $\eta=\eta_0$. However, it should be emphasized that due to the assumption of a very small Mach number [see Eq. (8) as well as Eq. (26)] the results derived in the following would remain unchanged even if the commonly used approximative power laws $\kappa=\kappa_0(\vartheta/\vartheta_0)^\beta$ and $\mu=\mu_0(\vartheta/\vartheta_0)^\beta$, where the coefficient $\beta=O(1)$, were adopted.

Since the fluid is presupposed to be a perfect gas, the equilibrium sound speed c_0 equals $(\gamma p_0/\rho_0)^{1/2}$ and $\vartheta_0=c_0^2/((\gamma-1)C_p)$. The two-dimensional Navier–Stokes equations in cylindrical coordinates for axisymmetric flow, the continuity equation, the energy equation, and the equation of state then read

$$\rho\frac{\partial v_z}{\partial t}+\rho v_z\frac{\partial v_z}{\partial z}+\frac{a}{\epsilon}\rho v_r\frac{\partial v_z}{\partial r}-\frac{\epsilon^2\alpha^2}{a^2d^2}\left(\frac{4}{3}+\eta\right)\frac{\partial^2 v_z}{\partial z^2}-\frac{\alpha^2}{d^2}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_z}{\partial r}\right)-\frac{\epsilon\alpha^2}{ad^2}\left(\frac{1}{3}+\eta\right)\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_r}{\partial z}\right)+\frac{\partial p}{\partial z}=0,\tag{17}$$

$$\rho\frac{\partial v_r}{\partial t}+\rho v_z\frac{\partial v_r}{\partial z}+\frac{a}{\epsilon}\rho v_r\frac{\partial v_r}{\partial r}-\frac{\epsilon^2\alpha^2}{a^2d^2}\frac{\partial^2 v_r}{\partial z^2}-\frac{\alpha^2}{d^2}\left(\frac{4}{3}+\eta\right)\times\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_r}{\partial r}\right)-\frac{v_r}{r^2}\right]-\frac{\epsilon\alpha^2}{ad^2}\left(\frac{1}{3}+\eta\right)\frac{\partial^2 v_z}{\partial z\partial r}+\frac{a}{\epsilon}\frac{\partial p}{\partial r}=0,\tag{18}$$

$$\frac{\partial\rho}{\partial t}+v_z\frac{\partial\rho}{\partial z}+\frac{a}{\epsilon}v_r\frac{\partial\rho}{\partial r}+\rho\frac{\partial v_z}{\partial z}+\frac{a}{\epsilon}\rho\frac{1}{r}\frac{\partial}{\partial r}(rv_r)=0,\tag{19}$$

$$\rho\frac{\partial\vartheta}{\partial t}+\rho v_z\frac{\partial\vartheta}{\partial z}+\frac{a}{\epsilon}\rho v_r\frac{\partial\vartheta}{\partial r}-(\gamma-1)\left(\frac{\partial p}{\partial t}+v_z\frac{\partial p}{\partial z}+\frac{a}{\epsilon}v_r\frac{\partial p}{\partial r}\right)-\frac{\epsilon^2\alpha^2}{a^2d^2\,\mathrm{Pr}}\frac{\partial^2\vartheta}{\partial z^2}-\frac{\alpha^2}{d^2\,\mathrm{Pr}}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\vartheta}{\partial r}\right)-\frac{1}{a^2d^2}(\gamma-1)\Phi=0,\tag{20}$$

$$\vartheta\rho=\gamma p.\tag{21}$$

Here, the superscripts $*$ introduced to indicate nondimensional quantities have already been omitted. The quantity Φ,

$$\Phi\sim\epsilon^2\alpha^2\times\max\left(\left(\frac{\partial v_z}{\partial z}\right)^2,\frac{a^2}{\epsilon^2}\left(\frac{\partial v_z}{\partial r}\right)^2,H\left(\frac{\partial v_r}{\partial z}\right)^2,\frac{a^2}{\epsilon^2}\left(\frac{\partial v_r}{\partial r}\right)^2,\frac{a}{\epsilon}\frac{\partial v_z}{\partial z}\frac{\partial v_r}{\partial r},\frac{a}{\epsilon}\frac{\partial v_z}{\partial r}\frac{\partial v_r}{\partial z}\right),\tag{22}$$

is the so-called dissipation function, which will turn out to be negligibly small as well.

Equations (17)–(21) will be solved subject to the boundary conditions at the tube wall

$$r=1:\quad v_z=v_r=0,\quad \vartheta=1,\tag{23}$$

and the symmetry conditions at the center of the tube

$$r=0:\quad v_r=\frac{\partial v_z}{\partial r}=\frac{\partial p}{\partial r}=\frac{\partial\vartheta}{\partial r}=0.\tag{24}$$

IV. APPLICATION OF THE MMS

Utilizing the parameters introduced in Eq. (16), the HFL can now simply be defined as

$$\epsilon\ll 1,\quad \alpha\ll 1\tag{25}$$

since then $\mathrm{Re}^{-1}\sim\epsilon^2\alpha^2\ll 1$, $l\sim\epsilon\ll 1$, and $\mathrm{St}^{-1}\sim\delta/R\sim\alpha\ll 1$, as required by conditions (5) and (11). In order to avoid nonlinear effects entering the first- and second-order terms, the Mach number is assumed to be of any order $O(\epsilon^i\alpha^j)$ such that

$$M\ll\epsilon^m\alpha^n,\quad m+n=2\tag{26}$$

holds, suggesting the use of M, α, and ϵ as perturbation parameters for an asymptotic analysis. Hence, the velocity components and the relevant thermodynamic quantities are expressed in the form:

$$
\begin{aligned}
v_z &= M(v_{z1} + \epsilon v_{z\epsilon} + \alpha v_{z\alpha} + \epsilon^2 v_{z\epsilon^2} + \epsilon\alpha v_{z\epsilon\alpha} + \alpha^2 v_{z\alpha^2} \\
&\quad + \cdots) + M^2(\cdots) + \cdots, \\
v_r &= M(v_{r1} + \epsilon v_{r\epsilon} + \alpha v_{r\alpha} + \epsilon^2 v_{r\epsilon^2} + \epsilon\alpha v_{r\epsilon\alpha} + \alpha^2 v_{r\alpha^2} \\
&\quad + \cdots) + M^2(\cdots) + \cdots, \\
p &= \frac{1}{\gamma} + M(p_1 + \epsilon p_\epsilon + \alpha p_\alpha + \epsilon^2 p_{\epsilon^2} + \epsilon\alpha p_{\epsilon\alpha} + \alpha^2 p_{\alpha^2} \\
&\quad + \cdots) + M^2(\cdots) + \cdots, \\
\rho &= 1 + M(\rho_1 + \epsilon\rho_\epsilon + \alpha\rho_\alpha + \epsilon^2\rho_{\epsilon^2} + \epsilon\alpha\rho_{\epsilon\alpha} + \alpha^2\rho_{\alpha^2} \\
&\quad + \cdots) + M^2(\cdots) + \cdots, \\
\vartheta &= 1 + M(\vartheta_1 + \epsilon\vartheta_\epsilon + \alpha\vartheta_\alpha + \epsilon^2\vartheta_{\epsilon^2} + \epsilon\alpha\vartheta_{\epsilon\alpha} + \alpha^2\vartheta_{\alpha^2} \\
&\quad + \cdots) + M^2(\cdots) + \cdots .
\end{aligned} \tag{27}
$$

The sound pressure is then given by $p_s = p - 1/\gamma$.

The investigation of the most significant features of acoustic waves emerging over spatial ranges in longitudinal direction of the orders $O(1)$ and $O(\alpha^{-1})$ requires at least the introduction of the length scales z and, additionally, $z_1 = \alpha z$. However, in order to resolve the changes of the second-order terms, e.g., $v_{z\alpha}$ and $v_{z\epsilon}$, over long distances as well, a third scale $z_2 = \alpha^2 z$ according to

$$
\frac{\partial}{\partial z} \to \frac{\partial}{\partial z} + \alpha\frac{\partial}{\partial z_1} + \alpha^2\frac{\partial}{\partial z_2} \tag{28}
$$

has to be used, because otherwise the generation of secular terms could not be avoided. Here it should be noted that the application of additional length scales proportional to powers of ϵ is not necessary since the solutions will retain their validity even for z as large as ϵ^{-1} or ϵ^{-2}. Results of a simplified study in the case of linear waves in tubes with a *rectangular* cross section that is restricted to the leading order terms and thus involves only the two length scales z and z_1 are presented in the Appendices A and B.

Substitution of the expressions (27) into the system (17)–(21) leads to a set of equations which is valid in the entire width of the tube except the boundary region and, therefore, is called the *outer* expansion. Close to the tube wall, where the stretched lateral coordinate

$$
s = \frac{1 - r}{\alpha} \tag{29}
$$

is of $O(1)$, viscosity and heat conduction play an important role; these effects have to be accounted for by a separate investigation of the boundary layer. Consequently, for the *inner* expansion, the coordinate r then has to be replaced with $1 - \alpha s$ and, furthermore,

$$
\frac{\partial}{\partial r} \to -\frac{1}{\alpha}\frac{\partial}{\partial s}. \tag{30}
$$

Please note that in the inner expansion, the density terms will be denoted by Ψ, whereas all other inner quantities will be written in capital letters, e.g., $V_{z\epsilon}$. In order to match the quantities arising from the two expansions, Van Dyke's *asymptotic matching principle* (see, e.g., Ref. 12, pp. 173–179) will be applied.

Since the resulting equations, more precisely, those that are relevant to solving the problem, will turn out to be linear, a Fourier transform with respect to the time could be used. Equivalently, each unsteady perturbation w or W is decomposed into a steady modal amplitude and a time-harmonic function such that

$$
\begin{aligned}
w(z, z_1, z_2, r, t) &= \Re(\hat{w}(z, z_1, z_2, r) e^{j\omega t}), \\
W(z, z_1, z_2, s, t) &= \Re(\hat{W}(z, z_1, z_2, s) e^{j\omega t}).
\end{aligned} \tag{31}
$$

Evaluation of the continuity equation (19) together with conditions (23) and (24) then yields

$$
\begin{aligned}
&\frac{\partial \hat{V}_{r1}}{\partial s} = \frac{\partial \hat{V}_{r\epsilon}}{\partial s} = \frac{\partial \hat{V}_{r\alpha}}{\partial s} = \frac{\partial \hat{V}_{r\epsilon^2}}{\partial s} = 0 \Rightarrow \hat{V}_{r1} = \hat{V}_{r\epsilon} = \hat{V}_{r\alpha} \\
&\qquad = \hat{V}_{r\epsilon^2} = 0,
\end{aligned} \tag{32}
$$

$$
j\omega\hat{\Psi}_1 + \frac{\partial \hat{V}_{z1}}{\partial z} - a\frac{\partial \hat{V}_{r\epsilon\alpha}}{\partial s} = 0, \tag{33}
$$

$$
j\omega\hat{\Psi}_\epsilon + \frac{\partial \hat{V}_{z\epsilon}}{\partial z} - a\frac{\partial \hat{V}_{r\epsilon^2\alpha}}{\partial s} = 0, \tag{34}
$$

$$
j\omega\hat{\Psi}_\alpha + \frac{\partial \hat{V}_{z\alpha}}{\partial z} + \frac{\partial \hat{V}_{z1}}{\partial z_1} - a\frac{\partial \hat{V}_{r\epsilon\alpha^2}}{\partial s} + a\hat{V}_{r\epsilon\alpha} = 0, \tag{35}
$$

and, furthermore,

$$
\frac{\partial}{\partial r}(r\hat{v}_{r1}) = \frac{\partial}{\partial r}(r\hat{v}_{r\alpha}) = 0 \Rightarrow \hat{v}_{r1} = \hat{v}_{r\alpha} = 0, \tag{36}
$$

$$
j\omega\hat{\rho}_1 + \frac{\partial \hat{v}_{z1}}{\partial z} + a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r\epsilon}) = 0, \tag{37}
$$

$$
j\omega\hat{\rho}_\epsilon + \frac{\partial \hat{v}_{z\epsilon}}{\partial z} + a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r\epsilon^2}) = 0, \tag{38}
$$

$$
j\omega\hat{\rho}_\alpha + \frac{\partial \hat{v}_{z\alpha}}{\partial z} + \frac{\partial \hat{v}_{z1}}{\partial z_1} + a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r\epsilon\alpha}) = 0, \tag{39}
$$

$$
j\omega\hat{\rho}_{\epsilon\alpha} + \frac{\partial \hat{v}_{z\epsilon\alpha}}{\partial z} + \frac{\partial \hat{v}_{z\epsilon}}{\partial z_1} + a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r\epsilon^2\alpha}) = 0, \tag{40}
$$

$$
j\omega\hat{\rho}_{\alpha^2} + \frac{\partial \hat{v}_{z\alpha^2}}{\partial z} + \frac{\partial \hat{v}_{z\alpha}}{\partial z_1} + \frac{\partial \hat{v}_{z1}}{\partial z_2} + a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r\epsilon\alpha^2}) = 0. \tag{41}
$$

As a consequence of Eqs. (32) and (36), the inner and the outer expansion of the Navier–Stokes equation (18) governing the radial motion give

$$
\frac{\partial \hat{P}_1}{\partial s} = \frac{\partial \hat{P}_\epsilon}{\partial s} = \frac{\partial \hat{P}_\alpha}{\partial s} = \frac{\partial \hat{P}_{\epsilon\alpha}}{\partial s} = \frac{\partial \hat{P}_{\alpha^2}}{\partial s} = 0, \tag{42}
$$

$$
\frac{\partial \hat{p}_1}{\partial r} = \frac{\partial \hat{p}_\epsilon}{\partial r} = \frac{\partial \hat{p}_\alpha}{\partial r} = \frac{\partial \hat{p}_{\epsilon\alpha}}{\partial r} = \frac{\partial \hat{p}_{\alpha^2}}{\partial r} = 0. \tag{43}
$$

This agrees with the expectation that the pressure in the boundary layer is set by the pressure fluctuations in the core region, which, due to the long wavelength assumption, are independent of the lateral coordinate r.

The set of equations following from the inner expansion of Eq. (21) reads

$$\hat{\Psi}_1=\gamma\hat{P}_1-\hat{\Theta}_1,\quad \hat{\Psi}_\epsilon=\gamma\hat{P}_\epsilon-\hat{\Theta}_\epsilon,\quad \hat{\Psi}_\alpha=\gamma\hat{P}_\alpha-\hat{\Theta}_\alpha,$$
$$\hat{\Psi}_{\epsilon\alpha}=\gamma\hat{P}_{\epsilon\alpha}-\hat{\Theta}_{\epsilon\alpha},\quad \hat{\Psi}_{\alpha^2}=\gamma\hat{P}_{\alpha^2}-\hat{\Theta}_{\alpha^2}, \tag{44}$$

which holds for the outer expansion as well.

Substituting the expansions (27) into the Navier–Stokes equation for the axial direction (17) and using Eqs. (32) and (36), one obtains for the inner quantities

$$j\omega\hat{V}_{z1}-\frac{1}{d^2}\frac{\partial^2\hat{V}_{z1}}{\partial s^2}+\frac{\partial\hat{P}_1}{\partial z}=0, \tag{45}$$

$$j\omega\hat{V}_{z\epsilon}-\frac{1}{d^2}\frac{\partial^2\hat{V}_{z\epsilon}}{\partial s^2}+\frac{\partial\hat{P}_\epsilon}{\partial z}=0, \tag{46}$$

$$j\omega\hat{V}_{z\alpha}-\frac{1}{d^2}\frac{\partial^2\hat{V}_{z\alpha}}{\partial s^2}+\frac{1}{d^2}\frac{\partial\hat{V}_{z1}}{\partial s}+\frac{\partial\hat{P}_\alpha}{\partial z}+\frac{\partial\hat{P}_1}{\partial z_1}=0, \tag{47}$$

and for the outer quantities

$$j\omega\hat{v}_{z1}+\frac{\partial\hat{p}_1}{\partial z}=0, \tag{48}$$

$$j\omega\hat{v}_{z\epsilon}+\frac{\partial\hat{p}_\epsilon}{\partial z}=0, \tag{49}$$

$$j\omega\hat{v}_{z\alpha}+\frac{\partial\hat{p}_\alpha}{\partial z}+\frac{\partial\hat{p}_1}{\partial z_1}=0, \tag{50}$$

$$j\omega\hat{v}_{z\epsilon\alpha}+\frac{\partial\hat{p}_{\epsilon\alpha}}{\partial z}+\frac{\partial\hat{p}_\epsilon}{\partial z_1}=0, \tag{51}$$

$$j\omega\hat{v}_{z\alpha^2}+\frac{\partial\hat{p}_{\alpha^2}}{\partial z}+\frac{\partial\hat{p}_\alpha}{\partial z_1}+\frac{\partial\hat{p}_1}{\partial z_2}=0, \tag{52}$$

where in the last Eq. (52) the relationship $\partial\hat{v}_{z1}/\partial r=0$ resulting from Eqs. (43) and (48) has already been applied; the other expansion terms of $\hat{v}_z$ appearing in Eqs. (49)–(52) also turn out to be independent of the radial coordinate r.

Hence, the dissipation function Φ defined in Eq. (22) is of the order $O(M^2\epsilon^2\alpha^2)$ in the core region and $O(M^2)$ in the boundary layer, and the expansions of the energy equation (21) take the form

$$j\omega\hat{\Theta}_1-j\omega(\gamma-1)\hat{P}_1-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_1}{\partial s^2}=0, \tag{53}$$

$$j\omega\hat{\Theta}_\epsilon-j\omega(\gamma-1)\hat{P}_\epsilon-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_\epsilon}{\partial s^2}=0, \tag{54}$$

$$j\omega\hat{\Theta}_\alpha-j\omega(\gamma-1)\hat{P}_\alpha-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_\alpha}{\partial s^2}+\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial\hat{\Theta}_1}{\partial s}=0, \tag{55}$$

$$\hat{\vartheta}_1=(\gamma-1)\hat{p}_1,\quad \hat{\vartheta}_\epsilon=(\gamma-1)\hat{p}_\epsilon,\quad \hat{\vartheta}_\alpha=(\gamma-1)\hat{p}_\alpha,$$
$$\hat{\vartheta}_{\epsilon\alpha}=(\gamma-1)\hat{p}_{\epsilon\alpha},\quad \hat{\vartheta}_{\alpha^2}=(\gamma-1)\hat{p}_{\alpha^2}. \tag{56}$$

Upon comparison with the expressions (44) this shows that

$$\hat{\rho}_1=\hat{p}_1,\quad \hat{\rho}_\epsilon=\hat{p}_\epsilon,\quad \hat{\rho}_\alpha=\hat{p}_\alpha,\quad \hat{\rho}_{\epsilon\alpha}=\hat{p}_{\epsilon\alpha},$$
$$\hat{\rho}_{\alpha^2}=\hat{p}_{\alpha^2}. \tag{57}$$

Solving systems (37)–(41) and (48)–(52) for the velocity components in radial direction subject to condition (24) gives

$$\hat{v}_{r\epsilon}=r\hat{f}_\epsilon(z,z_1,z_2),\quad \hat{v}_{r\epsilon^2}=r\hat{f}_{\epsilon^2}(z,z_1,z_2),$$
$$\hat{v}_{r\epsilon\alpha}=r\hat{f}_{\epsilon\alpha}(z,z_1,z_2), \tag{58}$$
$$\hat{v}_{r\epsilon^2\alpha}=r\hat{f}_{\epsilon^2\alpha}(z,z_1,z_2),\quad \hat{v}_{r\epsilon\alpha^2}=r\hat{f}_{\epsilon\alpha^2}(z,z_1,z_2).$$

Equation (32) in combination with Van Dyke's matching principle then implies that $\hat{f}_\epsilon=\hat{f}_{\epsilon^2}=0$. Thus, the wave equations resulting from Eqs. (37), (38), (48), and (49), which determine the evolution of the pressure fluctuations $\hat{p}_1=\hat{P}_1$, $\hat{p}_\epsilon=\hat{P}_\epsilon$ over distances of the order $O(1)$ (i.e., distances comparable to the wavelengths), are

$$\omega^2\hat{p}_1+\frac{\partial^2\hat{p}_1}{\partial z^2}=0,\quad \omega^2\hat{p}_\epsilon+\frac{\partial^2\hat{p}_\epsilon}{\partial z^2}=0. \tag{59}$$

Additionally, by applying the matching rules to the expressions for $\hat{V}_{z1}$ and $\hat{\Theta}_1$ derived from Eqs. (45), (53), and the boundary condition (23), one obtains

$$\hat{V}_{z1}=\hat{v}_{z1}(1-e^{-(1+j)ds\sqrt{(\omega/2)}}),$$
$$\hat{\Theta}_1=\hat{p}_1(\gamma-1)(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}}). \tag{60}$$

Precisely the same functional dependence on the boundary layer coordinate s is valid for $\hat{V}_{z\epsilon}$ and $\hat{\Theta}_\epsilon$ as well. As a consequence, Eqs. (33) and (34) can be solved to give

$$\hat{V}_{r\epsilon\alpha}=\hat{p}_1\frac{1+j}{ad}\sqrt{\frac{\omega}{2}}\left[\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})\right.$$
$$\left.+(1-e^{-(1+j)ds\sqrt{(\omega/2)}})\right],$$
$$\hat{V}_{r\epsilon^2\alpha}=\hat{p}_\epsilon\frac{1+j}{ad}\sqrt{\frac{\omega}{2}}\left[\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})\right. \tag{61}$$
$$\left.+(1-e^{-(1+j)ds\sqrt{(\omega/2)}})\right].$$

Carrying out the matching procedure with the expansion terms of $\hat{v}_r$ given by Eq. (58) leads to

$$\hat{f}_{\epsilon\alpha}(z,z_1,z_2)=\hat{p}_1\frac{j\omega}{a}\hat{F},\quad \hat{f}_{\epsilon^2\alpha}(z,z_1,z_2)=\hat{p}_\epsilon\frac{j\omega}{a}\hat{F}, \tag{62}$$

where

$$\hat{F}=\frac{1-j}{d\sqrt{2\omega}}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right), \tag{63}$$

and it then follows from Eqs. (39), (40), (50), and (51) that

$$\omega^2\hat{p}_\alpha+\frac{\partial^2\hat{p}_\alpha}{\partial z^2}=-2\frac{\partial^2\hat{p}_1}{\partial z\partial z_1}-2\omega^2\hat{p}_1\hat{F},$$
$$\omega^2\hat{p}_{\epsilon\alpha}+\frac{\partial^2\hat{p}_{\epsilon\alpha}}{\partial z^2}=-2\frac{\partial^2\hat{p}_\epsilon}{\partial z\partial z_1}-2\omega^2\hat{p}_\epsilon\hat{F}. \tag{64}$$

Inspection of the relationships (64) shows that in order to rule out secular solutions of the form $\hat{p}_\alpha\propto z$, $\hat{p}_{\epsilon\alpha}\propto z$, the resonant forcing terms on the right-hand sides must be annihilated, generating the wave equations

$$\omega^2\hat{p}_\alpha+\frac{\partial^2\hat{p}_\alpha}{\partial z^2}=0,\quad \omega^2\hat{p}_{\epsilon\alpha}+\frac{\partial^2\hat{p}_{\epsilon\alpha}}{\partial z^2}=0,$$
$$\frac{\partial^2\hat{p}_1}{\partial z\partial z_1}+\omega^2\hat{p}_1\hat{F}=0,\quad \frac{\partial^2\hat{p}_\epsilon}{\partial z\partial z_1}+\omega^2\hat{p}_\epsilon\hat{F}=0, \tag{65}$$

which govern the propagation of the pressure perturbations $\hat{p}_\alpha$ and $\hat{p}_{\epsilon\alpha}$ over distances of order $O(1)$ and, respectively, the propagation of the pressure perturbations $\hat{p}_1$ and $\hat{p}_\epsilon$ over distances of order $O(\alpha^{-1})$. The imaginary part $\Im(-\omega\hat{F})$ is thus identified as the leading order decay rate of the sound pressure due to boundary layer attenuation.[1] Worth mentioning is the fact that the equation for $\hat{p}_1$ represents the multiple scales equivalent to the model equation that was derived by Pierce[18] (pp. 531–534) from a variational principle.

Proceeding in very much the same way as before, Eqs. (47) and (55) are solved for the axial velocity and the temperature in the boundary layer:

$$\hat{V}_{z\alpha}=\hat{v}_{z\alpha}(1-e^{-(1+j)ds\sqrt{(\omega/2)}})-\hat{v}_{z1}\frac{s}{2}e^{-(1+j)ds\sqrt{(\omega/2)}},$$
$$\hat{\Theta}_\alpha=\hat{p}_\alpha(\gamma-1)(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})-\hat{p}_1(\gamma-1)\frac{s}{2}e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}}. \tag{66}$$

Substituting these expressions together with the relationships for $\hat{\Psi}_\alpha$, $\hat{V}_{z1}$, and $\hat{V}_{r\epsilon\alpha}$ given in Eqs. (44), (60), and (61), respectively, into Eq. (35), the following result for the inner quantity $\hat{V}_{r\epsilon\alpha^2}$ can be derived:

$$\hat{V}_{r\epsilon\alpha^2}=\hat{p}_\alpha\frac{1+j}{ad}\sqrt{\frac{\omega}{2}}\left[\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})+(1-e^{-(1+j)ds\sqrt{(\omega/2)}})\right]-\hat{p}_1\frac{\gamma-1}{2ad^2\,\mathrm{Pr}}\times\left[1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}}+(1+j)ds\sqrt{\frac{\omega\,\mathrm{Pr}}{2}}\times e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}}\right]-\hat{p}_1\frac{1}{2ad^2}\left[1-e^{-(1+j)ds\sqrt{(\omega/2)}}+(1+j)ds\sqrt{\frac{\omega}{2}}e^{-(1+j)ds\sqrt{(\omega/2)}}\right]+\hat{p}_1\frac{1+j}{ad}\sqrt{2\omega}\hat{F}\times(1-e^{-(1+j)ds\sqrt{(\omega/2)}})-\hat{p}_1\frac{j\omega}{a}\hat{F}s. \tag{67}$$

When the matching procedure is again applied to the inner and outer radial velocity components, one then finds that

$$\hat{f}_{\epsilon\alpha^2}(z,z_1,z_2)=\frac{j\omega}{a}(\hat{p}_\alpha\hat{F}+\hat{p}_1\hat{G}), \tag{68}$$

with

$$\hat{G}=\frac{j}{2d^2\omega}\left(1+\frac{\gamma-1}{\mathrm{Pr}}\right)+\frac{1-j}{d}\sqrt{\frac{2}{\omega}}\hat{F}=-\frac{j}{2d^2\omega}\left(3-\frac{\gamma-1}{\mathrm{Pr}}+4\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right). \tag{69}$$

After substituting $\hat{v}_{r\epsilon\alpha^2}$ from Eqs. (58) and (68) into Eq. (41), Eq. (52) can be recast into

$$\omega^2\hat{p}_{\alpha^2}+\frac{\partial^2\hat{p}_{\alpha^2}}{\partial z^2}=-2\frac{\partial^2\hat{p}_1}{\partial z\partial z_2}-2\frac{\partial^2\hat{p}_\alpha}{\partial z\partial z_1}+\omega^2(\hat{p}_1\hat{F}^2-2\hat{p}_1\hat{G}-2\hat{p}_\alpha\hat{F}), \tag{70}$$

where the right-hand side is identified as resonant forcing, since it would involve secular terms in $\hat{p}_{\alpha^2}$. This implies that the quantity $\hat{p}_\alpha$ has to satisfy the solvability condition

$$\frac{\partial^2\hat{p}_\alpha}{\partial z\partial z_1}+\omega^2\hat{p}_\alpha\hat{F}=-\frac{\partial^2\hat{p}_1}{\partial z\partial z_2}-\omega^2\hat{p}_1\left(\hat{G}-\frac{\hat{F}^2}{2}\right). \tag{71}$$

In turn, unless this right-hand side is annihilated, it would inevitably lead to $\hat{p}_\alpha$ being proportional to z_1. Thus, the resulting equations read

$$\omega^2\hat{p}_{\alpha^2}+\frac{\partial^2\hat{p}_{\alpha^2}}{\partial z^2}=0,\quad \frac{\partial^2\hat{p}_\alpha}{\partial z\partial z_1}+\omega^2\hat{p}_\alpha\hat{F}=0,$$
$$\frac{\partial^2\hat{p}_1}{\partial z\partial z_2}+\omega^2\hat{p}_1\hat{H}=0. \tag{72}$$

Here

$$\hat{H}=\hat{G}-\frac{\hat{F}^2}{2}=-\frac{j}{d^2\omega}\left[1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left(1-\frac{\gamma}{2\sqrt{\mathrm{Pr}}}\right)\right], \tag{73}$$

with $\Im(-\omega\hat{H})$ being the correction term to the attenuation rate $\Im(-\omega\hat{F})$.

The solutions of the wave equations (59), (65), and (72) can then be written as

$$\hat{p}_1=\hat{c}_{11}e^{-j\omega(z+\hat{F}z_1+\hat{H}z_2)}+\hat{c}_{21}e^{j\omega(z+\hat{F}z_1+\hat{H}z_2)},$$
$$\hat{p}_\epsilon=\hat{c}_{1\epsilon}(z_2)e^{-j\omega(z+\hat{F}z_1)}+\hat{c}_{2\epsilon}(z_2)e^{j\omega(z+\hat{F}z_1)},$$
$$\hat{p}_\alpha=\hat{c}_{1\alpha}(z_2)e^{-j\omega(z+\hat{F}z_1)}+\hat{c}_{2\alpha}(z_2)e^{j\omega(z+\hat{F}z_1)},$$
$$\hat{p}_{\epsilon\alpha}=\hat{c}_{1\epsilon\alpha}(z_1,z_2)e^{-j\omega z}+\hat{c}_{2\epsilon\alpha}(z_1,z_2)e^{j\omega z},$$
$$\hat{p}_{\alpha^2}=\hat{c}_{1\alpha^2}(z_1,z_2)e^{-j\omega z}+\hat{c}_{2\alpha^2}(z_1,z_2)e^{j\omega z}. \tag{74}$$

A. Results

The goal pursued in the study presented here is to derive the asymptotically correct expressions for the series impedance $\hat{Z}$ and the shunt admittance $\hat{Y}$, which were introduced in Eq. (1), or, equivalently, the expressions for tube param-

eters $\hat{Z}_c$ and $\hat{\Gamma}$ defined in Eq. (3). In order to calculate the volume flow $\hat{u}$, the axial Navier–Stokes equation (17) is averaged over the cross-sectional area and expanded with respect to the perturbation parameters ϵ and α, giving

$$
\begin{aligned}
&j\omega\hat{\bar{v}}_{z1}+\frac{\partial\hat{p}_1}{\partial z}=0,\\
&j\omega\hat{\bar{v}}_{z\epsilon}+\frac{\partial\hat{p}_\epsilon}{\partial z}=0,\\
&j\omega\hat{\bar{v}}_{z\alpha}+\frac{2}{d^2}\frac{\partial\hat{V}_{z1}}{\partial s}\bigg|_{s=0}+\frac{\partial\hat{p}_\alpha}{\partial z}+\frac{\partial\hat{p}_1}{\partial z_1}=0,\\
&j\omega\hat{\bar{v}}_{z\epsilon\alpha}+\frac{2}{d^2}\frac{\partial\hat{V}_{z\epsilon}}{\partial s}\bigg|_{s=0}+\frac{\partial\hat{p}_{\epsilon\alpha}}{\partial z}+\frac{\partial\hat{p}_\epsilon}{\partial z_1}=0,\\
&j\omega\hat{\bar{v}}_{z\alpha^2}+\frac{2}{d^2}\frac{\partial\hat{V}_{z\alpha}}{\partial s}\bigg|_{s=0}+\frac{\partial\hat{p}_{\alpha^2}}{\partial z}+\frac{\partial\hat{p}_\alpha}{\partial z_1}+\frac{\partial\hat{p}_1}{\partial z_2}=0.
\end{aligned}
\tag{75}
$$

Furthermore, as shown in Eq. (28), the multiple scales technique applied here requires the derivatives with respect to z appearing in the definitions of $\hat{Z}$ and $\hat{Y}$ to be replaced with derivatives with respect to the three length scales z, z_1, and z_2. Using the wave equations (59), (65), and (72) governing the sound pressure and the expansion terms of the volume flow $\hat{u}=\hat{\bar{v}}_z$ that can be deduced from Eq. (75), the following expressions for the series impedance and the shunt admittance are obtained:

$$
\begin{aligned}
\hat{Z}&=-\frac{1}{\hat{u}}\left(\frac{d}{dz}+\alpha\frac{d}{dz_1}+\alpha^2\frac{d}{dz_2}\right)\hat{p}_s\\
&=j\omega+\alpha\frac{1+j}{d}\sqrt{2\omega}+\alpha^2\frac{3}{d^2}+O(\epsilon\alpha^2,\alpha^3),\\
\hat{Y}&=-\frac{1}{\hat{p}_s}\left(\frac{d}{dz}+\alpha\frac{d}{dz_1}+\alpha^2\frac{d}{dz_2}\right)\hat{u}\\
&=j\omega+\alpha\frac{1+j}{d}(\gamma-1)\sqrt{\frac{2\omega}{\mathrm{Pr}}}-\alpha^2\frac{\gamma-1}{d^2\,\mathrm{Pr}}+O(\epsilon\alpha^2,\alpha^3).
\end{aligned}
\tag{76}
$$

The expansions for tube parameters $\hat{Z}_c$ and $\hat{\Gamma}$ then read

$$
\begin{aligned}
\hat{Z}_c&=1+\alpha\frac{1-j}{d\sqrt{2\omega}}\left(1-\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)-\alpha^2\frac{j}{2d^2\omega}\left(2-2\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right.\\
&\quad\left.-\frac{5\gamma-3\gamma^2-2}{\mathrm{Pr}}\right)+\cdots,\\
\hat{\Gamma}&=j\omega+\alpha j\omega\hat{F}+\alpha^2 j\omega\hat{H}+\cdots\\
&=j\omega+\alpha\frac{1+j}{d}\sqrt{\frac{\omega}{2}}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)+\alpha^2\frac{1}{d^2}\left[1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right.\\
&\quad\left.\times\left(1-\frac{\gamma}{2\sqrt{\mathrm{Pr}}}\right)\right]+\cdots.
\end{aligned}
\tag{77}
$$

It should be emphasized that the expression for the characteristic impedance $\hat{Z}_c$ could have been truncated after the second term as well, since the investigation is confined to the first- and second-order terms of the sound pressure and the volume flow. However, as far as the propagation parameter is concerned, all three expansion terms have to be included, because the length of the tube might be much larger than the wavelengths, i.e., even of the order $O(\alpha^{-1})$.

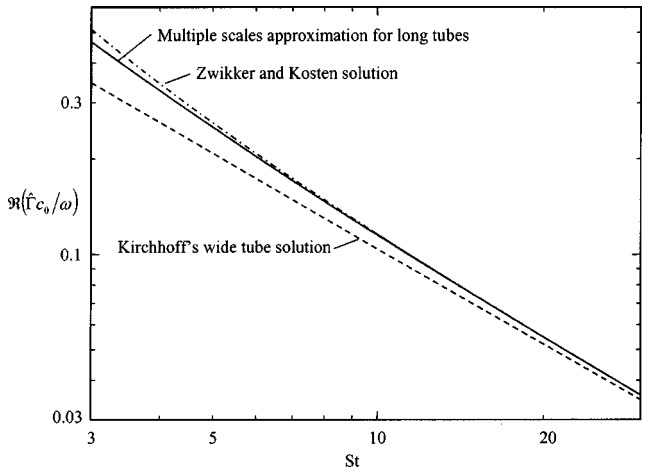


FIG. 2. Graphs of $\Re(\hat{\Gamma}c_0/\omega)$ as functions of St in double logarithmic scale; Pr=0.707, γ=1.402 (air at 300 K); St$\gg$1: $\Re(\hat{\Gamma}c_0/\omega)=O(\mathrm{St}^{-1})$.

The corresponding results for the dimensional transmission line parameters are presented in Appendix C. Interestingly, Eq. (C2) completely conforms to the results given by Keefe,[10] which were derived by high order series expansions of the Zwikker and Kosten solutions (9) with respect to the (small) inverse of the Stokes number. Comparison with the expressions (14) clearly shows that the terms of the order $O(\mathrm{St}^{-2})$ appearing in Eq. (C2) must not be neglected if the effects emerging over distances $L\gg\lambda$ are to be incorporated into the HFL of the series impedance and the shunt admittance. In addition to providing precise solutions for the sound pressure, the velocity components, and the other thermodynamic quantities, the study presented here thus extends the validity of the Zwikker and Kosten approach to the case of sound propagation in long tubes in the limit of large Stokes numbers (HFL), provided that the low reduced frequency assumptions (5) hold. In order to demonstrate this equivalence graphically, the different solutions for the real part of the propagation parameter resulting from Eqs. (9), (14), and (C2) are depicted in Fig. 2.

B. Example

The results summarized so far are sufficient to evaluate the transmission line parameters entering the four-pole transfer matrix $\hat{\mathbf{A}}$ defined in Eq. (2) for a long circular duct with isothermal walls. Exemplarily, the total load impedance $\hat{Z}_t=\hat{p}_{s-}/\hat{u}_-$ of a tube radiating into open space will be calculated. The parameters α and ϵ are required to be small, in order to comply with the requirements the multiple scales analysis elaborated in Sec. IV is based on. Furthermore, the tube length L shall be so large that the nondimensional quantity

$$
L_1=\alpha L \tag{78}
$$

is of order $O(1)$. These conditions can be satisfied if, e.g., a thin tube of dimensional length 10 cm and radius 0.5 mm is considered, the fluid is assumed to be air (at 300 K) and the characteristic dimensional wavelength λ equals 7 mm, which corresponds to a driving frequency f of approximately 50 000 Hz and leads to $\epsilon/a=0.071\ll 1$, $\alpha/d=0.036\ll 1$, $L=14.286\gg 1$. Then one may choose $\epsilon=\alpha=0.071$, $a=1$, and $d=2.008$, resulting in $L_1=1.020=O(1)$.

Since the condition $\epsilon\ll 1$ holds, the radiation impedance can be expanded into

$$\hat{Z}_e=\epsilon\hat{Z}_{e\epsilon}+\epsilon^2\hat{Z}_{e\epsilon^2}+\cdots. \tag{79}$$

In the case of radiation of sound from a circular duct with an infinite flange, the leading order term $\hat{Z}_{e\epsilon}$ is a pure imaginary and given by

$$\hat{Z}_{e\epsilon}=L_e j\omega, \tag{80}$$

where $L_e=0.8217/a$ is the so-called quasistatic end correction. The validity of this result will be proved in Sec. IV C.

Substituting the solutions from Eq. (77) and the relationship $\hat{p}_{s+}/\hat{u}_+=\hat{Z}_e$ into Eq. (4), a series expansion of the resulting total impedance with respect to ϵ and α then yields (see Appendix C for the dimensional form of this result)

$$\begin{aligned}\hat{Z}_t=\frac{\hat{p}_{s-}}{\hat{u}_-}&=\tanh\left(j\omega L_1\left(\frac{1}{\alpha}+\hat{F}\right)\right)+\epsilon\frac{\hat{Z}_{e\epsilon}}{\cosh\left(j\omega L_1\left(\frac{1}{\alpha}+\hat{F}\right)\right)^2}\\&+\alpha\left[\frac{1-j}{d\sqrt{2\omega}}\left(1-\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\tanh\left(j\omega L_1\left(\frac{1}{\alpha}+\hat{F}\right)\right)\right.\\&\left.+\frac{j\omega L_1\hat{H}}{\cosh\left(j\omega L_1\left(\frac{1}{\alpha}+\hat{F}\right)\right)^2}\right]+\cdots\\&=\tanh\left(L_1\left[\frac{j\omega}{\alpha}+\frac{1+j}{d}\sqrt{\frac{\omega}{2}}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\right]\right)\\&+\alpha\left[\frac{1-j}{d\sqrt{2\omega}}\left(1-\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\right.\\&\left.\times\tanh\left(L_1\left[\frac{j\omega}{\alpha}+\frac{1+j}{d}\sqrt{\frac{\omega}{2}}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\right]\right)\right.\\&\left.+\frac{\frac{\epsilon}{\alpha}\hat{Z}_{e\epsilon}+\frac{L_1}{d^2}\left[1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left(1-\frac{\gamma}{2\sqrt{\mathrm{Pr}}}\right)\right]}{\cosh\left(L_1\left[\frac{j\omega}{\alpha}+\frac{1+j}{d}\sqrt{\frac{\omega}{2}}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\right]\right)^2}\right]+\cdots,\end{aligned} \tag{81}$$

which reveals the fact that in the case of $\epsilon\sim\alpha$, the effects resulting from the radiation at the end of the tube enter the expansion of the total impedance at the same order as the viscothermal effects accumulating over the considerably large length of the tube. It should be emphasized that the arguments of the functions $1/\cosh(\cdot)$ and $\tanh(\cdot)$ always contain a real part of order $O(1)$, since L_1 is assumed to be of order $O(1)$. Thus, the moduli of these terms will remain $O(1)$ quantities as well, even if the dimensional tube length assumes a value close to $(n+1/2)c_0/(2f)$, where n is an integer of $O(\alpha^{-1})$, i.e., $L_1\approx\alpha(n+1/2)\pi/\omega$. This is in contrast to the well-known resonance phenomenon occurring in shorter tubes, where thermal and viscous effects contribute much less to the total load impedance.

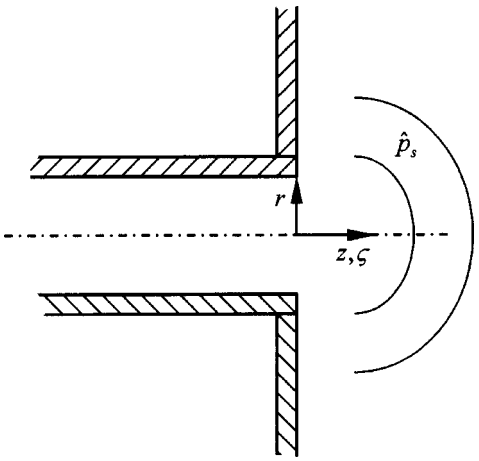


FIG. 3. Sketch and notation of a duct radiating into half-space.

C. End correction for circular tubes with an infinite flange

The radiation impedance at the flanged opening will excite higher-order modes in the backward propagating wave. However, due to the long wavelength assumption $l\sim\epsilon\ll 1$ these modes have cut-off frequencies well above the driving frequency $\omega/(2\pi)$ and, consequently, die out rapidly within a spatial range comparable to the radius of the cross section, leaving only the lowest mode to propagate over any longer distance along the tube (see Ref. 14, p. 499). In order to resolve the details of the flow close to the exit of the duct, a separate perturbation analysis is therefore necessary: For convenience, the origin of the axial coordinate is set to the position of the opening as shown in Fig. 3. Similarly to the inner expansion for the boundary layer, a stretched inner coordinate for the end region

$$\zeta=\frac{z}{\epsilon},\quad \frac{\partial}{\partial z}\to\frac{1}{\epsilon}\frac{\partial}{\partial\zeta} \tag{82}$$

can then be introduced. The quantities arising from this second inner expansion will be denoted by the subscript e.

Substitution of the expressions (27) into the basic equations (17)–(21) in the same manner as was done in Sec. IV leads to

$$\frac{\partial\hat{V}_{re1}}{\partial s}=0\Rightarrow\hat{V}_{re1}=0, \tag{83}$$

$$\frac{\partial\hat{v}_{ze1}}{\partial\zeta}+a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{re1})=0, \tag{84}$$

$$\frac{\partial \hat{P}_{e1}}{\partial s}=\frac{\partial \hat{P}_{e\epsilon}}{\partial s}=\frac{\partial \hat{P}_{e\alpha}}{\partial s}=0, \tag{85}$$

$$\frac{\partial \hat{p}_{e1}}{\partial r}=\frac{\partial \hat{p}_{e\alpha}}{\partial r}=0, \quad j\omega\hat{v}_{re1}+a\frac{\partial \hat{p}_{e\epsilon}}{\partial r}=0, \tag{86}$$

$$\hat{\Psi}_{e1}=\gamma\hat{P}_{e1}-\hat{\Theta}_{e1}, \quad \hat{\Psi}_{e\epsilon}=\gamma\hat{P}_{e\epsilon}-\hat{\Theta}_{e\epsilon},$$

$$\hat{\Psi}_{e\alpha}=\gamma\hat{P}_{e\alpha}-\hat{\Theta}_{e\alpha}, \tag{87}$$

$$\hat{\rho}_{e1}=\gamma\hat{p}_{e1}-\hat{\vartheta}_{e1}, \quad \hat{\rho}_{e\epsilon}=\gamma\hat{p}_{e\epsilon}-\hat{\vartheta}_{e\epsilon},$$

$$\hat{\rho}_{e\alpha}=\gamma\hat{p}_{e\alpha}-\hat{\vartheta}_{e\alpha}, \tag{88}$$

$$\frac{\partial \hat{P}_{e1}}{\partial \zeta}=\frac{\partial \hat{P}_{e\alpha}}{\partial \zeta}=0, \quad j\omega\hat{V}_{ze1}-\frac{1}{d^2}\frac{\partial^2 \hat{V}_{ze1}}{\partial s^2}+\frac{\partial \hat{P}_{e\epsilon}}{\partial \zeta}=0, \tag{89}$$

$$\frac{\partial \hat{p}_{e1}}{\partial \zeta}=\frac{\partial \hat{p}_{e\alpha}}{\partial \zeta}=0, \quad j\omega\hat{v}_{ze1}+\frac{\partial \hat{p}_{e\epsilon}}{\partial \zeta}=0, \tag{90}$$

$$j\omega\hat{\Theta}_{e1}-j\omega(\gamma-1)\hat{P}_{e1}-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_{e1}}{\partial s^2}=0,$$

$$j\omega\hat{\Theta}_{e\epsilon}-j\omega(\gamma-1)\hat{P}_{e\epsilon}-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_{e\epsilon}}{\partial s^2}=0, \tag{91}$$

$$j\omega\hat{\Theta}_{e\alpha}-j\omega(\gamma-1)\hat{P}_{e\alpha}-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_{e\alpha}}{\partial s^2}=0,$$

$$\hat{\vartheta}_{e1}=(\gamma-1)\hat{p}_{e1}, \quad \hat{\vartheta}_{e\epsilon}=(\gamma-1)\hat{p}_{e\epsilon},$$

$$\hat{\vartheta}_{e\alpha}=(\gamma-1)\hat{p}_{e\alpha}, \tag{92}$$

showing that the leading order term of the sound pressure and the correction term $\hat{p}_{e\alpha}$ do not change at all in the end region $\zeta\leqslant 0$. Obviously, the second-order term $\hat{p}_{e\epsilon}$ satisfies the Laplace equation

$$\zeta\leqslant 0: \quad \frac{\partial^2\hat{p}_{e\epsilon}}{\partial\zeta^2}+a^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\hat{p}_{e\epsilon}}{\partial r}\right)=0, \tag{93}$$

subject to the boundary condition

$$r=1: \quad \frac{\partial\hat{p}_{e\epsilon}}{\partial r}=0. \tag{94}$$

In the region outside the tube where the dimensional axial and the dimensional radial coordinate are comparable to the wavelength, i.e., $z=O(1)$ and $r=O(1/\epsilon)$, the expansion terms of the pressure fluctuations are determined by a set of Helmholtz equations. However, close to the mouth where $\zeta=O(1)$ and $r=O(1)$, the Helmholtz equation for $\hat{p}_{e\epsilon}$ reduces to the Laplace equation so that Eq. (93) turns out to hold equally well for $\zeta>0$. Moreover, the quantities $\hat{p}_{e1}$ and $\hat{p}_{e\alpha}$ are found to fulfill Eqs. (86) and (90) outside the tube too. Since the acoustic waves spread hemispherically, the pressure perturbation $\hat{p}_{e\epsilon}$ in the region $\zeta>0$ thus is related to the axial velocity in the opening by

$$\zeta>0: \quad \hat{p}_{e\epsilon}(\zeta,r)=j\omega\frac{1}{2a\pi}\int_0^1 r_+\hat{v}_{ze1}(r_+)\int_0^{2\pi}\frac{1}{\sqrt{\zeta^2+r^2+r_+^2-2rr_+\cos(\beta)}}d\beta dr_+$$

$$=j\omega\frac{1}{a}\int_0^1 r_+\hat{v}_{ze1}(r_+)\int_0^\infty e^{-\zeta\tau}\mathrm{J}_0(\tau r)\mathrm{J}_0(\tau r_+)d\tau dr_+, \tag{95}$$

with $\hat{v}_{ze1}(r_+)$ given by the relationship in Eq. (90), whereas one obtains for the other expansion terms

$$\zeta>0: \quad \hat{p}_{e1}(\zeta,r)=0, \quad \hat{p}_{e\alpha}(\zeta,r)=0. \tag{96}$$

Evaluation of Eqs. (86) and (90) renders Eq. (96) valid for $\zeta\leqslant 0$ as well. As a consequence, Van Dyke's matching rules applied to the outer solutions (74) give

$$\hat{c}_{11}=-\hat{c}_{21}, \quad \hat{c}_{1\alpha}(0)=-\hat{c}_{2\alpha}(0) \tag{97}$$

and, furthermore,

$$\zeta\to-\infty: \quad \hat{p}_{e\epsilon}\to\hat{c}_{1\epsilon}(0)+\hat{c}_{2\epsilon}(0)-2j\omega\hat{c}_{11}\zeta. \tag{98}$$

Interestingly, Eqs. (93), (94), (98), together with the boundary condition that can be deduced from Eq. (95) by taking the limit as $\zeta\to 0+$, turn out to constitute precisely the same system of equations that was solved by Rayleigh[2] (pp. 487–491, see also Ref. 19) in order to calculate the so-called quasistatic end correction

$$L_e=\frac{\hat{c}_{1\epsilon}(0)+\hat{c}_{2\epsilon}(0)}{2j\omega\hat{c}_{11}}. \tag{99}$$

A summary of this investigation can also be found in the paper by Howe.[20] Using a variational approach based on trial functions, Rayleigh obtained an approximate value of the end correction as $L_e=0.8242/a$. In addition, Daniell[21] provided a solution bounded by the narrow range $0.82141<aL_e<0.82168$. More recently, other authors, e.g., Norris and Sheng,[22] calculated the quasistatic as well as the dynamic reflection of sound from the end of a flanged pipe by implementing a rational function approximation with the Bessel functions used as basis functions. In these studies, more accurate numerical results for the end correction in the limit of zero frequency are presented which are all very close to the value $L_e=0.8217/a$ also presented in Sec. IV B.

Finally, introducing the relationship (99) in combination with condition (97) into the expressions for the sound pressure (74) and the volume flow (75) and performing a series

expansion of $\hat{Z}_e=\hat{p}_{s+}/\hat{u}_+$ with respect to the perturbation parameters ϵ and α yields the result given by Eq. (80).

V. EXTENSION: THE CASE $l=O(1)$

The following part will deal with the propagation of sound waves in cylindrical ducts, proceeding from the assumption that the characteristic wavelength is comparable to the diameter of the tube. In this case the HFL is defined as

$$\epsilon=1, \quad \alpha\ll 1, \tag{100}$$

leading to $\mathrm{Re}^{-1}\sim\alpha^2\ll 1$, $l\sim\mathrm{He}=O(1)$, and $\mathrm{St}^{-1}\sim\delta/R\sim\alpha\ll 1$.

As a consequence, the inner and outer expansions of the basic equations (17)–(21) subject to conditions (23) and (24) are carried out with respect to the remaining perturbation parameter α only. One then obtains for the continuity equation:

$$\frac{\partial\hat{V}_{r1}}{\partial s}=0\Rightarrow\hat{V}_{r1}=0, \tag{101}$$

$$j\omega\hat{\Psi}_1+\frac{\partial\hat{V}_{z1}}{\partial z}-a\frac{\partial\hat{V}_{r\alpha}}{\partial s}=0, \tag{102}$$

$$j\omega\hat{\Psi}_\alpha+\frac{\partial\hat{V}_{z\alpha}}{\partial z}+\frac{\partial\hat{V}_{z1}}{\partial z_1}-a\frac{\partial\hat{V}_{r\alpha^2}}{\partial s}+a\hat{V}_{r\alpha}=0, \tag{103}$$

$$j\omega\hat{\rho}_1+\frac{\partial\hat{v}_{z1}}{\partial z}+a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r1})=0, \tag{104}$$

$$j\omega\hat{\rho}_\alpha+\frac{\partial\hat{v}_{z\alpha}}{\partial z}+\frac{\partial\hat{v}_{z1}}{\partial z_1}+a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r\alpha})=0, \tag{105}$$

$$j\omega\hat{\rho}_{\alpha^2}+\frac{\partial\hat{v}_{z\alpha^2}}{\partial z}+\frac{\partial\hat{v}_{z\alpha}}{\partial z_1}+\frac{\partial\hat{v}_{z1}}{\partial z_2}+a\frac{1}{r}\frac{\partial}{\partial r}(r\hat{v}_{r\alpha^2})=0, \tag{106}$$

for the radial Navier–Stokes equation:

$$\frac{\partial\hat{P}_1}{\partial s}=\frac{\partial\hat{P}_\alpha}{\partial s}=0, \tag{107}$$

$$j\omega\hat{V}_{r\alpha}-a\frac{\partial\hat{P}_{\alpha^2}}{\partial s}=0, \tag{108}$$

$$j\omega\hat{v}_{r1}+a\frac{\partial\hat{p}_1}{\partial r}=0, \quad j\omega\hat{v}_{r\alpha}+a\frac{\partial\hat{p}_\alpha}{\partial r}=0, \tag{109}$$

$$j\omega\hat{v}_{r\alpha^2}-\frac{1}{a^2d^2}\frac{\partial^2v_{r1}}{\partial z^2}-\frac{1}{d^2}\left(\frac{4}{3}+\eta\right)\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_{r1}}{\partial r}\right)-\frac{v_{r1}}{r^2}\right]-\frac{1}{ad^2}\left(\frac{1}{3}+\eta\right)\frac{\partial^2v_{z1}}{\partial z\partial r}+a\frac{\partial\hat{p}_{\alpha^2}}{\partial r}=0, \tag{110}$$

for the equation of state:

$$\hat{\Psi}_1=\gamma\hat{P}_1-\hat{\Theta}_1, \quad \hat{\Psi}_\alpha=\gamma\hat{P}_\alpha-\hat{\Theta}_\alpha, \quad \hat{\Psi}_{\alpha^2}=\gamma\hat{P}_{\alpha^2}-\hat{\Theta}_{\alpha^2}, \tag{111}$$

which holds for the outer expansion as well, furthermore, for the axial Navier–Stokes equation:

$$j\omega\hat{V}_{z1}-\frac{1}{d^2}\frac{\partial^2\hat{V}_{z1}}{\partial s^2}+\frac{\partial\hat{P}_1}{\partial z}=0, \tag{112}$$

$$j\omega\hat{V}_{z\alpha}-\frac{1}{d^2}\frac{\partial^2\hat{V}_{z\alpha}}{\partial s^2}+\frac{1}{d^2}\frac{\partial\hat{V}_{z1}}{\partial s}+\frac{\partial\hat{P}_\alpha}{\partial z}+\frac{\partial\hat{P}_1}{\partial z_1}=0, \tag{113}$$

$$j\omega\hat{v}_{z1}+\frac{\partial\hat{p}_1}{\partial z}=0, \tag{114}$$

$$j\omega\hat{v}_{z\alpha}+\frac{\partial\hat{p}_\alpha}{\partial z}+\frac{\partial\hat{p}_1}{\partial z_1}=0, \tag{115}$$

$$j\omega\hat{v}_{z\alpha^2}-\frac{1}{a^2d^2}\left(\frac{4}{3}+\eta\right)\frac{\partial^2v_{z1}}{\partial z^2}-\frac{1}{d^2}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_{z1}}{\partial r}\right)-\frac{1}{ad^2}\left(\frac{1}{3}+\eta\right)\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_{r1}}{\partial z}\right)+\frac{\partial\hat{p}_{\alpha^2}}{\partial z}+\frac{\partial\hat{p}_\alpha}{\partial z_1}+\frac{\partial\hat{p}_1}{\partial z_2}=0, \tag{116}$$

and, finally, for the energy equation:

$$j\omega\hat{\Theta}_1-j\omega(\gamma-1)\hat{P}_1-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_1}{\partial s^2}=0, \tag{117}$$

$$j\omega\hat{\Theta}_\alpha-j\omega(\gamma-1)\hat{P}_\alpha-\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\Theta}_\alpha}{\partial s^2}+\frac{1}{d^2\,\mathrm{Pr}}\frac{\partial\hat{\Theta}_1}{\partial s}=0, \tag{118}$$

$$\hat{\vartheta}_1=(\gamma-1)\hat{p}_1, \quad \hat{\vartheta}_\alpha=(\gamma-1)\hat{p}_\alpha, \tag{119}$$

$$j\omega\hat{\vartheta}_{\alpha^2}-j\omega(\gamma-1)\hat{p}_{\alpha^2}-\frac{1}{a^2d^2\,\mathrm{Pr}}\frac{\partial^2\hat{\vartheta}_1}{\partial z^2}-\frac{1}{d^2\,\mathrm{Pr}}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\hat{\vartheta}_1}{\partial r}\right)=0. \tag{120}$$

Substitution of Eq. (111) into Eq. (119) yields

$$\hat{\rho}_1=\hat{p}_1, \quad \hat{\rho}_\alpha=\hat{p}_\alpha \tag{121}$$

Inspection of Eqs. (101), (104), (109), and (114) shows that the leading order pressure perturbations in the core region $\hat{p}_1$ satisfy the Helmholtz equation

$$\omega^2\hat{p}_1+\frac{\partial^2\hat{p}_1}{\partial z^2}+a^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\hat{p}_1}{\partial r}\right)=0, \tag{122}$$

together with the boundary condition

$$r=1: \quad \frac{\partial\hat{p}_1}{\partial r}=0. \tag{123}$$

The solution is given by the modal decomposition

$$\hat{p}_1=\sum_{n=0}^{\infty}\left(\hat{C}_{1n1}(z_1,z_2)\frac{\mathrm{J}_0(\gamma_nr)}{\mathrm{J}_0(\gamma_n)}e^{-j\hat{k}_nz}+\hat{C}_{2n1}(z_1,z_2)\frac{\mathrm{J}_0(\gamma_nr)}{\mathrm{J}_0(\gamma_n)}e^{j\hat{k}_nz}\right), \tag{124}$$

where the parameters γ_n, $n=0,1,2,...$ are the zeros of the first-order Bessel function $J_1(\hat{\xi})$, with $\gamma_0=0$ and the quantities

$$\hat{k}_n=\begin{cases}\sqrt{\omega^2-a^2\gamma_n^2}, & \omega>a\gamma_n\\ -j\sqrt{a^2\gamma_n^2-\omega^2}, & \omega<a\gamma_n\end{cases} \tag{125}$$

are, respectively, the axial wave numbers for the propagating and evanescent modes.

Combining Eqs. (105), (109), (114), and (115) yields the inhomogeneous Helmholtz equation for the second-order term $\hat{p}_\alpha$,

$$\omega^2\hat{p}_\alpha+\frac{\partial^2\hat{p}_\alpha}{\partial z^2}+a^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\hat{p}_\alpha}{\partial r}\right)=-2\frac{\partial^2\hat{p}_1}{\partial z\,\partial z_1}. \tag{126}$$

With Van Dyke's matching rules applied to the expressions for $\hat{V}_{z1}$, $\hat{v}_{z1}$, $\hat{\Theta}_1$, and $\hat{\vartheta}_1$ resulting from Eqs. (112), (114), (117), (119), and (23), the solutions

$$\begin{aligned}&\hat{V}_{z1}=\hat{v}_{z1}|_{r=1}(1-e^{-(1+j)ds\sqrt{(\omega/2)}}),\\ &\hat{\Theta}_1=\hat{p}_1|_{r=1}(\gamma-1)(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})\end{aligned} \tag{127}$$

are derived. It then follows from Eqs. (102) and (111) that

$$\begin{aligned}a\hat{V}_{r\alpha}=\hat{p}_1|_{r=1}&\left[j\omega s+\frac{1+j}{d}\sqrt{\frac{\omega}{2}}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right.\\ &\left.\times(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})\right]+\frac{\partial^2\hat{p}_1}{\partial z^2}\Bigg|_{r=1}\\ &\times\frac{1}{\omega^2}\left[j\omega s-\frac{1+j}{d}\sqrt{\frac{\omega}{2}}(1-e^{-(1+j)ds\sqrt{(\omega/2)}})\right].\end{aligned} \tag{128}$$

Consequently, the matching principle gives the boundary condition for the outer expansion term $\hat{p}_\alpha$ in the form

$$r=1:\quad \frac{\partial\hat{p}_\alpha}{\partial r}=\frac{1-j}{a^2d}\omega\sqrt{\frac{\omega}{2}}\left(\hat{p}_1\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}-\frac{\partial^2\hat{p}_1}{\partial z^2}\frac{1}{\omega^2}\right). \tag{129}$$

The solution $\hat{p}'_\alpha$ to the homogeneous part of Eq. (126) subject to the homogeneous boundary conditions at $r=1$ assumes the same form as the expression for $\hat{p}_1$ from Eq. (124). However, the functions $\hat{C}_{1n1}$ and $\hat{C}_{2n1}$ have to be replaced with $\hat{C}_{1n\alpha}$ and $\hat{C}_{2n\alpha}$, respectively. In order to determine a particular solution $\hat{p}''_\alpha$ such that $\hat{p}_\alpha=\hat{p}'_\alpha+\hat{p}''_\alpha$ the ansatz

$$\begin{aligned}\hat{p}''_\alpha=\sum_{n=0}^{\infty}&\left[\left(\hat{f}_{1n\alpha}(z,z_1,z_2)\frac{J_0(\gamma_n r)}{J_0(\gamma_n)}\right.\right.\\ &\left.+\hat{g}_{n\alpha}(r)\hat{C}_{1n1}(z_1,z_2)\right)e^{-j\hat{k}_n z}\\ &+\left(\hat{f}_{2n\alpha}(z,z_1,z_2)\frac{J_0(\gamma_n r)}{J_0(\gamma_n)}\right.\\ &\left.\left.+\hat{g}_{n\alpha}(r)\hat{C}_{2n1}(z_1,z_2)\right)e^{j\hat{k}_n z}\right]\end{aligned} \tag{130}$$

is made, which is required to fulfill the conditions (24) and (129). Furthermore, the functional dependence on the coordinate r of the terms that are generated by the left-hand side of Eq. (126) for each mode $n=0,1,2,...$ on substituting Eq. (130) must be given by the radial eigenfunctions $J_0(\gamma_n r)$. Together, these requirements imply that

$$\hat{g}_{0\alpha}=\frac{\omega^2\hat{F}_0}{2a^2}r^2,\quad \hat{g}_{n\alpha}=\frac{\hat{k}_n^2\hat{F}_n}{a^2}\frac{rJ_1(\gamma_n r)}{\gamma_n J_0(\gamma_n)},\quad n=1,2,..., \tag{131}$$

with the parameters $\hat{F}_n$ defined as

$$\hat{F}_n=\frac{1-j}{d\sqrt{2\omega}}\left(1+\frac{\omega^2}{\hat{k}_n^2}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right),\quad n=0,1,2,.... \tag{132}$$

Substitution of expression (130) into Eq. (126) then results in

$$\begin{aligned}&\frac{\partial^2\hat{f}_{1n\alpha}}{\partial z^2}-2j\hat{k}_n\frac{\partial\hat{f}_{1n\alpha}}{\partial z}=-2\hat{F}_n\hat{k}_n^2\hat{C}_{1n1}+2j\hat{k}_n\frac{\partial\hat{C}_{1n1}}{\partial z_1},\\ &\frac{\partial^2\hat{f}_{2n\alpha}}{\partial z^2}+2j\hat{k}_n\frac{\partial\hat{f}_{2n\alpha}}{\partial z}=-2\hat{F}_n\hat{k}_n^2\hat{C}_{2n1}-2j\hat{k}_n\frac{\partial\hat{C}_{2n1}}{\partial z_1}.\end{aligned} \tag{133}$$

The forcing terms on the right-hand sides are resonant and would produce secular terms in in the functions $\hat{f}_{1n\alpha}$ and $\hat{f}_{2n\alpha}$. Therefore, they must be annihilated, yielding

$$\begin{aligned}&\hat{C}_{1n1}(z_1,z_2)\to\hat{C}_{1n1}(z_2)e^{-j\hat{k}_n\hat{F}_n z_1},\\ &\hat{C}_{2n1}(z_1,z_2)\to\hat{C}_{2n1}(z_2)e^{j\hat{k}_n\hat{F}_n z_1},\end{aligned} \tag{134}$$

whereupon the functions $\hat{f}_{1n\alpha}$ and $\hat{f}_{2n\alpha}$ can simply be set to zero, since any other solution would lead to expressions for $\hat{p}''_\alpha$ that could be incorporated into the homogeneous solution $\hat{p}'_\alpha$. Please note that Eq. (134) could also have been derived by using a different concept: As a result of the homogeneous problem

$$(\omega^2-\hat{k}_n^2)\hat{\psi}+a^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\hat{\psi}}{\partial r}\right)=0,\quad r=1:\quad \frac{\partial\hat{\psi}}{\partial r}=0 \tag{135}$$

having a nontrivial solution and the operator being self-adjoint, the inhomogeneous problem has a solution only if the forcing terms are orthogonal to the homogeneous solution, see, e.g., Ref. 11. This is known as the *Fredholm alternative* and introduces two solvability conditions for Eq. (126), which are identical to the right-hand sides of Eq. (133).

The quantities $\Im(-\omega\hat{F}_0)=\Im(-\omega\hat{F})$ and $\Im(-\hat{k}_n\hat{F}_n)$, $n=1,2,...$, turn out to be, respectively, the leading order decay rate of the fundamental mode and the leading order decay rates of the higher order modes due to boundary layer attenuation, which is in accordance with Beatty's[23] results for the axisymmetric case obtained by using the concept of boundary layer admittance.

In order to calculate the corresponding expressions for the third-order terms, a procedure very similar to that used

before is performed: Combination of Eqs. (106), (110), (114)–(116), (119), and (120) and multiple application of Eq. (122) shows that

$$\hat{v}_{z\alpha^2}=-\frac{1}{a^2d^2}\frac{\partial\hat{p}_1}{\partial z}\left(\frac{4}{3}+\eta\right)+\frac{j}{\omega}\left(\frac{\partial\hat{p}_1}{\partial z_2}+\frac{\partial\hat{p}_\alpha}{\partial z_1}+\frac{\partial\hat{p}_{\alpha^2}}{\partial z}\right) \tag{136}$$

and that $\hat{p}_{\alpha^2}$ has to satify the inhomogeneous equation

$$\begin{aligned}\omega^2\hat{p}_{\alpha^2}+\frac{\partial^2\hat{p}_{\alpha^2}}{\partial z^2}+a^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\hat{p}_{\alpha^2}}{\partial r}\right)&\\ =-2\frac{\partial^2\hat{p}_1}{\partial z\partial z_2}-\frac{\partial^2\hat{p}_1}{\partial z_1^2}+j\omega^3\frac{1}{a^2d^2}\hat{p}_1\left(\frac{4}{3}+\eta+\frac{\gamma-1}{\mathrm{Pr}}\right)&\\ -2\frac{\partial^2\hat{p}_\alpha}{\partial z\partial z_1},&\end{aligned} \tag{137}$$

where the third term on the right-hand side obviously incorporates the effects of heat conduction and viscosity in the core region associated with the propagation of the leading order pressure fluctuations $\hat{p}_1$. This is in contrast to the analysis in Sec. IV where, due to the long wavelength assumption $l\ll 1$, these effects do not affect the wave equations [see Eq. (70)]. Moreover, Eqs. (113) and (118) are solved for the axial velocity and the temperature in the boundary layer, leading to

$$\begin{aligned}\hat{V}_{z\alpha}&=\hat{v}_{z\alpha}|_{r=1}(1-e^{-(1+j)ds\sqrt{(\omega/2)}})\\ &\quad-\hat{v}_{z1}|_{r=1}\frac{s}{2}e^{-(1+j)ds\sqrt{(\omega/2)}},\\ \hat{\Theta}_\alpha&=\hat{p}_\alpha|_{r=1}(\gamma-1)(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})\\ &\quad-\hat{p}_1|_{r=1}(\gamma-1)\frac{s}{2}e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}}.\end{aligned} \tag{138}$$

After substitution for $\hat{\Psi}_\alpha$, $\hat{V}_{z1}$, $\hat{V}_{r\alpha}$, and $\hat{V}_{z\alpha}$ from Eqs. (111), (127), (128), and (138), respectively, integration of Eq. (103) with respect to the inner coordinate s gives

$$\begin{aligned}a\hat{V}_{r\alpha^2}=\hat{p}_1|_{r=1}&\left[\frac{j\omega}{2}s^2-\frac{\gamma-1}{2d^2\mathrm{Pr}}(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})+\frac{1+j}{2d}\sqrt{\frac{\omega}{2}}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}s(2-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})\right]\\ &+\left.\frac{\partial^2\hat{p}_1}{\partial z^2}\right|_{r=1}\frac{1}{\omega^2}\left[\frac{j\omega}{2}s^2+\frac{1}{2d^2}(1-e^{-(1+j)ds\sqrt{(\omega/2)}})-\frac{1+j}{2d}\sqrt{\frac{\omega}{2}}s(2-e^{-(1+j)ds\sqrt{(\omega/2)}})\right]\\ &+\hat{p}_\alpha|_{r=1}\left[j\omega s+\frac{1+j}{d}\sqrt{\frac{\omega}{2}}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}(1-e^{-(1+j)ds\sqrt{(\omega\,\mathrm{Pr}/2)}})\right]\\ &+\left(\left.\frac{\partial^2\hat{p}_\alpha}{\partial z^2}\right|_{r=1}+2\left.\frac{\partial^2\hat{p}_1}{\partial z\partial z_1}\right|_{r=1}\right)\frac{1}{\omega^2}\left[j\omega s-\frac{1+j}{d}\sqrt{\frac{\omega}{2}}(1-e^{-(1+j)ds\sqrt{(\omega/2)}})\right].\end{aligned} \tag{139}$$

Furthermore, upon applying the matching rules to the inner and outer expansions of the radial velocity, the boundary condition

$$\begin{aligned}r=1:\quad\frac{\partial\hat{p}_{\alpha^2}}{\partial r}&=\frac{j\omega}{2a^2d^2}\left(\hat{p}_1\frac{\gamma-1}{\mathrm{Pr}}-\frac{\partial^2\hat{p}_1}{\partial z^2}\frac{1}{\omega^2}\right)\\ &+\frac{1-j}{a^2d}\omega\sqrt{\frac{\omega}{2}}\left[\hat{p}_\alpha\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}-\left(\frac{\partial^2\hat{p}_\alpha}{\partial z^2}+2\frac{\partial^2\hat{p}_1}{\partial z\partial z_1}\right)\frac{1}{\omega^2}\right]\end{aligned} \tag{140}$$

is obtained. It should be mentioned that the particular solution $\hat{p}''_\alpha$ vanishes at the boundary for all higher order modes $n=1,2,\ldots$.

A particular solution $\hat{p}''_{\alpha^2}$ to Eq. (137) can be found by employing the ansatz

$$\begin{aligned}\hat{p}''_{\alpha^2}=\sum_{n=0}^{\infty}\Bigg[&\left(\hat{f}_{1n\alpha^2}(z,z_1,z_2)\frac{\mathrm{J}_0(\gamma_n r)}{\mathrm{J}_0(\gamma_n)}\right.\\ &+\hat{g}_{n\alpha^2}(r)\hat{C}_{1n1}(z_2)e^{-j\hat{F}_n\hat{k}_nz_1}\\ &\left.+\hat{h}_{n\alpha^2}(r)\hat{C}_{1n\alpha}(z_1,z_2)\right)e^{-j\hat{k}_nz}\\ &+\left(\hat{f}_{2n\alpha^2}(z,z_1,z_2)\frac{\mathrm{J}_0(\gamma_n r)}{\mathrm{J}_0(\gamma_n)}\right.\\ &+\hat{g}_{n\alpha^2}(r)\hat{C}_{2n1}(z_2)e^{j\hat{F}_n\hat{k}_nz_1}\\ &\left.+\hat{h}_{n\alpha^2}(r)\hat{C}_{2n\alpha}(z_1,z_2)\right)e^{j\hat{k}_nz}\Bigg],\end{aligned} \tag{141}$$

such that conditions (24) and (140) are satisfied and the operators on the left-hand side of Eq. (137) produce terms that have the same functional dependence on r as the terms on the right-hand side. As a consequence, one obtains

$$\hat{g}_{0\alpha^2}=\frac{\omega^2}{2a^2}\left(\hat{G}_0+\frac{\omega^2\hat{F}_0^2}{4a^2}\right)r^2+\frac{\omega^4\hat{F}_0^2}{16a^4}r^4,$$

$$\hat{g}_{n\alpha^2}=\frac{\hat{k}_n^2}{a^2}\left(\hat{G}_n+\frac{\hat{k}_n^2\hat{F}_n^2}{a^2\gamma_n^2}\right)\frac{rJ_1(\gamma_n r)}{\gamma_n J_0(\gamma_n)}$$

$$-\frac{\hat{k}_n^4\hat{F}_n^2}{2a^4}\frac{r^2J_0(\gamma_n r)}{\gamma_n^2J_0(\gamma_n)},\quad n=1,2,..., \tag{142}$$

$$\hat{h}_{0\alpha^2}=\frac{\omega^2\hat{F}_0}{2a^2}r^2,\quad \hat{h}_{n\alpha^2}=\frac{\hat{k}_n^2\hat{F}_n}{a^2}\frac{rJ_1(\gamma_n r)}{\gamma_n J_0(\gamma_n)},\quad n=1,2,...,$$

where

$$\hat{G}_n=-\frac{j}{2d^2\omega}\left[3-\frac{\omega^2}{\hat{k}_n^2}(\gamma-1)\left(\frac{1}{\mathrm{Pr}}-\frac{4}{\sqrt{\mathrm{Pr}}}\right)\right],$$

$$n=0,1,2,\ldots. \tag{143}$$

Substitution of Eq. (130) into Eq. (137) then yields

$$\frac{\partial^2\hat{f}_{1n\alpha^2}}{\partial z^2}-2j\hat{k}_n\frac{\partial\hat{f}_{1n\alpha^2}}{\partial z}=\left(-2\hat{H}_n\hat{k}_n^2\hat{C}_{1n1}+2j\hat{k}_n\frac{\partial\hat{C}_{1n1}}{\partial z_2}\right)$$

$$\times e^{-j\hat{F}_n\hat{k}_n z_1}-2\hat{F}_n\hat{k}_n^2\hat{C}_{1n\alpha}$$

$$+2j\hat{k}_n\frac{\partial\hat{C}_{1n\alpha}}{\partial z_1}, \tag{144}$$

$$\frac{\partial^2\hat{f}_{2n\alpha^2}}{\partial z^2}+2j\hat{k}_n\frac{\partial\hat{f}_{2n\alpha^2}}{\partial z}=\left(-2\hat{H}_n\hat{k}_n^2\hat{C}_{1n1}-2j\hat{k}_n\frac{\partial\hat{C}_{2n1}}{\partial z_2}\right)$$

$$\times e^{-j\hat{F}_n\hat{K}_n z_1}-2\hat{F}_n\hat{k}_n^2\hat{C}_{2n\alpha}$$

$$-2j\hat{k}_n\frac{\partial\hat{C}_{2n\alpha}}{\partial z_1},$$

where

$$\hat{H}_0=\hat{G}_0-\frac{\hat{F}_0^2}{2}+\frac{\omega^2\hat{F}_0^2}{4a^2}-\frac{j\omega}{2a^2d^2}\left(\frac{4}{3}+\eta+\frac{\gamma-1}{\mathrm{Pr}}\right)$$

$$=-\frac{j}{d^2\omega}\left[1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left(1-\frac{\gamma}{2\sqrt{\mathrm{Pr}}}\right)\right]$$

$$-\frac{j\omega}{2a^2d^2}\left[\frac{11}{6}+\eta+\frac{\gamma-1}{2\sqrt{\mathrm{Pr}}}\left(2+\frac{1+\gamma}{\sqrt{\mathrm{Pr}}}\right)\right],$$

$$\hat{H}_n=\hat{G}_n-\frac{\hat{F}_n^2}{2}-\frac{j\omega^3}{2a^2d^2\hat{k}_n^2}\left(\frac{4}{3}+\eta+\frac{\gamma-1}{\mathrm{Pr}}\right)$$

$$=-\frac{j}{d^2\omega}\left[1+\frac{\omega^2}{\hat{k}_n^2}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left[1-\frac{1}{2\sqrt{\mathrm{Pr}}}\left(1+\frac{\omega^2}{\hat{k}_n^2}(\gamma-1)\right)\right]\right]$$

$$-\frac{j\omega^3}{2a^2d^2\hat{k}_n^2}\left(\frac{4}{3}+\eta+\frac{\gamma-1}{\mathrm{Pr}}\right),\quad n=1,2,\ldots. \tag{145}$$

As before, the right-hand sides of Eq. (144) have to be annihilated in order to rule out secular terms in the functions $\hat{f}_{1n\alpha^2}$ and $\hat{f}_{2n\alpha^2}$. The resulting solvability conditions thus read

$$2\hat{F}_n\hat{k}_n^2\hat{C}_{1n\alpha}-2j\hat{k}_n\frac{\partial\hat{C}_{1n\alpha}}{\partial z_1}=\left(-2\hat{H}_n\hat{k}_n^2\hat{C}_{1n1}+2j\hat{k}_n\frac{\partial\hat{C}_{1n1}}{\partial z_2}\right)$$

$$\times e^{-j\hat{F}_n\hat{k}_n z_1},$$

$$2\hat{F}_n\hat{k}_n^2\hat{C}_{2n\alpha}+2j\hat{k}_n\frac{\partial\hat{C}_{2n\alpha}}{\partial z_1}=\left(-2\hat{H}_n\hat{k}_n^2\hat{C}_{1n1}-2j\hat{k}_n\frac{\partial\hat{C}_{2n1}}{\partial z_2}\right) \tag{146}$$

$$\times e^{-j\hat{F}_n\hat{k}_n z_1},$$

,

and, furthermore, $\hat{f}_{1n\alpha^2}$ and $\hat{f}_{2n\alpha^2}$ can be set to zero. The right-hand sides of Eq. (146) are again identified as resonant forcing terms, finally leading to

$$\hat{C}_{1n1}(z_2)e^{-j\hat{F}_n\hat{k}_n z_1}\rightarrow\hat{C}_{1n1}e^{-j\hat{k}_n(\hat{F}_n z_1+\hat{H}_n z_2)},$$

$$\hat{C}_{2n1}(z_2)e^{j\hat{F}_n\hat{k}_n z_1}\rightarrow\hat{C}_{2n1}e^{j\hat{k}_n(\hat{F}_n z_1+\hat{H}_n z_2)} \tag{147}$$

and, additionally,

$$\hat{C}_{1n\alpha}(z_1,z_2)\rightarrow\hat{C}_{1n\alpha}(z_2)e^{-j\hat{k}_n\hat{F}_n z_1},$$

$$\hat{C}_{2n\alpha}(z_1,z_2)\rightarrow\hat{C}_{2n\alpha}(z_2)e^{j\hat{k}_n\hat{F}_n z_1}. \tag{148}$$

Results. One interesting property of the parameter $\hat{H}_0$ for the fundamental mode is that it cannot be derived from the quantities $\hat{H}_n$, $n=1,2,...$, for the higher order modes simply by setting $\hat{k}_n=\omega$. However, it reduces to the parameter $\hat{H}$ defined in Sec. IV, when the limit as $a\rightarrow\infty$ is taken and thus, formally, the long wavelength assumption $l\ll 1$ is reintroduced. In accordance with this, the axial wave numbers $\hat{k}_n$, $n=1,2,...$, then approach $-j\infty$ and all higher order modes die away immediately.

The definition of the propagation parameter from Eq. (3) can easily be extended, such that each mode is treated separately: To this end, the quantities

$$\hat{\Gamma}_n=\sqrt{\frac{1}{\hat{p}_{ns}\hat{v}_{nz}}\frac{\partial\hat{p}_{ns}}{\partial z}\frac{\partial\hat{v}_{nz}}{\partial z}}=\sqrt{\frac{1}{\hat{p}_{ns}}\frac{\partial^2\hat{p}_{ns}}{\partial z^2}}, \tag{149}$$

where $\hat{p}_{ns}$ and $\hat{v}_{nz}$ denote the sound pressures and the axial velocities associated with the modes $n=0,1,2,...$, are introduced. Upon replacing the partial derivatives with respect to z with derivatives with respect to the three length scales used here according to Eq. (28), application of Eqs. (134), (147), and (148) implicates that the parameters $\hat{\Gamma}_n$ are given by the relationship

$$\hat{\Gamma}_n=\sqrt{-\hat{k}_n^2[1+2\alpha\hat{F}_n+\alpha^2(2\hat{H}_n+\hat{F}_n^2)]+\cdots},$$

$$n=0,1,2,..., \tag{150}$$

where $\hat{F}_n$ and $\hat{H}_n$ are the quantities already introduced in Eqs. (132) and (145), respectively. If $|\hat{k}_n|=O(1)$ this result simplifies to

$$\hat{\Gamma}_n=j\hat{k}_n+\alpha j\hat{k}_n\hat{F}_n+\alpha^2 j\hat{k}_n\hat{H}_n+\ldots,\quad n=0,1,2,\ldots. \tag{151}$$

Inspection of the definitions (132) and (145) shows that in the limiting case of a mode $n=m$ having a cut-off frequency so close to the driving frequency that $|\hat{k}_m| \ll 1$ the orders of magnitude of the different terms appearing in Eq. (150) may change completely. Before turning to a detailed analysis of that problem, please note that one aspect of the first- and second-order solutions derived by Anderson and Vaidya[13] in their study of linear sound propagation in slit-shaped waveguides carries over unchanged to the case of circular cross sections considered here: The solvability conditions resulting from Eqs. (133) and (144) are singular at the cut-off frequencies $\omega = a\gamma_n$, $n=1,2,...$, predicting that in the limit as $\omega \to a\gamma_m$, $\hat{k}_m \to 0$, the attenuation of the mode $n=m$ takes place over a much shorter spatial range than that defined by $z_1 = O(1)$. A closer examination of the solvability conditions then indicates that in order to investigate the modulation of a mode m featuring a cut-off frequency in proximity to the driving frequency ω such that

$$\omega = \omega_m + \alpha\omega_\alpha, \quad \omega_m = a\gamma_m, \quad m \geqslant 1 \tag{152}$$

holds, for this mth mode, the length scales z, z_1, and z_2 have to be replaced with the length scales

$$\overline{z_1} = \alpha^{1/2} z, \quad \overline{z_2} = \alpha^{3/2} z. \tag{153}$$

Moreover, this necessitates expressing the pressure and, similarly, the velocity components and the other thermodynamic quantities in the form:

$$p = \frac{1}{\gamma} + M(p_1 + \alpha^{1/2} p_{\alpha^{1/2}} + \alpha p_\alpha + \alpha^{3/2} p_{\alpha^{3/2}} + \alpha^2 p_{\alpha^2} + \cdots) + M^2(\cdots) + \cdots. \tag{154}$$

Application of the MMS then leads to the following results: (a) The leading order solutions given by Eqs. (109), (114), and (124) remain unchanged. However, the functions $\hat{C}_{1m1}(z_1, z_2)$ and $\hat{C}_{2m1}(z_1, z_2)$ have to be replaced with $\hat{C}_{1m1}(\overline{z_1}, \overline{z_2})$ and $\hat{C}_{2m1}(\overline{z_1}, \overline{z_2})$, respectively, and, additionally, ω must be replaced with ω_m, resulting in $\hat{k}_m = 0$. (b) The perturbations $\hat{v}_{r\alpha^{1/2}}$, $\hat{v}_{z\alpha^{1/2}}$, and $\hat{p}_{\alpha^{1/2}}$ satisfy Eq. (109),

$$j\omega_m \hat{v}_{z\alpha^{1/2}} + \frac{\partial \hat{p}_{\alpha^{1/2}}}{\partial z} + \frac{\partial \hat{p}_1}{\partial \overline{z_1}} = 0, \tag{155}$$

and the homogeneous Helmholz equation (122) subject to the boundary condition (123). (c) The sound propagation at the order $O(\alpha)$ is now governed by the set

$$\begin{aligned}
&j\omega_m \hat{v}_{r\alpha} + j\omega_\alpha \hat{v}_{r1} + a\frac{\partial \hat{p}_\alpha}{\partial r} = 0,\\
&j\omega_m \hat{v}_{z\alpha} + j\omega_\alpha \hat{v}_{z1} + \frac{\partial \hat{p}_\alpha}{\partial z} + \frac{\partial \hat{p}_{\alpha^{1/2}}}{\partial \overline{z_1}} + \frac{\partial \hat{p}_1}{\partial z_1} = 0,\\
&\omega_m^2 \hat{p}_\alpha + \frac{\partial^2 \hat{p}_\alpha}{\partial z^2} + a^2 \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \hat{p}_\alpha}{\partial r}\right)\\
&\quad = -2\omega_m \omega_\alpha \hat{p}_1 - 2\frac{\partial^2 \hat{p}_1}{\partial z \partial z_1} - \frac{\partial^2 \hat{p}_1}{\partial \overline{z_1}^2},
\end{aligned} \tag{156}$$

and the boundary condition (129). By introducing the ansatz (130), with $\hat{g}_{m\alpha}$ redefined as

$$\hat{g}_{m\alpha} = \frac{\hat{\bar{F}}_m}{a^2}\frac{rJ_1(\gamma_m r)}{\gamma_m J_0(\gamma_m)}, \quad \hat{\bar{F}}_m = \omega_m^{3/2}\frac{1-j}{d\sqrt{2}}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}, \tag{157}$$

the following modified solvability conditions are obtained:

$$\begin{aligned}
&-\hat{C}_{1n1}(\hat{F}_n \hat{k}_n^2 + \omega_m\omega_\alpha) + j\hat{k}_n \frac{\partial \hat{C}_{1n1}}{\partial z_1} + j\hat{k}_n \frac{\partial \hat{C}_{1n\alpha^{1/2}}}{\partial \overline{z_1}} = 0,\\
&\hat{C}_{2n1}(\hat{F}_n \hat{k}_n^2 + \omega_m\omega_\alpha) + j\hat{k}_n \frac{\partial \hat{C}_{2n1}}{\partial z_1} + j\hat{k}_n \frac{\partial \hat{C}_{2n\alpha^{1/2}}}{\partial \overline{z_1}} = 0,\\
&\quad n \neq m,
\end{aligned} \tag{158}$$

where ω in $\hat{k}_n$ and $\hat{F}_n$ has to be replaced with ω_m, and

$$2(\hat{C}_{1m1} + \hat{C}_{2m1})(\hat{\bar{F}}_m + \omega_m\omega_\alpha) + \frac{\partial^2}{\partial \overline{z_1}^2}(\hat{C}_{1m1} + \hat{C}_{2m1}) = 0 \tag{159}$$

for the mth mode. Equations (158) and (159) thus yield

$$\begin{aligned}
&\hat{C}_{1n1}(z_1, z_2) \to \hat{C}_{1n1}(z_2) e^{-j\hat{k}_n(\hat{F}_n + \omega_m\omega_\alpha/\hat{k}_n^2)z_1},\\
&\hat{C}_{2n1}(z_1, z_2) \to \hat{C}_{2n1}(z_2) e^{j\hat{k}_n(\hat{F}_n + \omega_m\omega_\alpha/\hat{k}_n^2)z_1}, \quad n \neq m,
\end{aligned} \tag{160}$$

$$\begin{aligned}
&\hat{C}_{1m1}(\overline{z_1}, \overline{z_2}) \to \hat{C}_{1m1}(\overline{z_2}) e^{-j(2\hat{\bar{F}}_m + 2\omega_m\omega_\alpha)^{1/2}\overline{z_1}},\\
&\hat{C}_{2m1}(\overline{z_1}, \overline{z_2}) \to \hat{C}_{2m1}(\overline{z_2}) e^{j(2\hat{\bar{F}}_m + 2\omega_m\omega_\alpha)^{1/2}\overline{z_1}}.
\end{aligned} \tag{161}$$

(d) Furthermore, the solvability conditions for $\hat{p}_{\alpha^{3/2}}$ imply that

$$\begin{aligned}
&\hat{C}_{1m\alpha^{1/2}}(\overline{z_1}, \overline{z_2}) \to \hat{C}_{1m\alpha^{1/2}}(\overline{z_2}) e^{-j(2\hat{\bar{F}}_m + 2\omega_m\omega_\alpha)^{1/2}\overline{z_1}},\\
&\hat{C}_{2m\alpha^{1/2}}(\overline{z_1}, \overline{z_2}) \to \hat{C}_{2m\alpha^{1/2}}(\overline{z_2}) e^{j(2\hat{\bar{F}}_m + 2\omega_m\omega_\alpha)^{1/2}\overline{z_1}}.
\end{aligned} \tag{162}$$

Substitution of ω from Eq. (152) into Eq. (134) multiplied by $\exp(\mp j\hat{k}_n z)$, series expansion with respect to α, and comparison with Eq. (160) immediately shows the equivalence of both formulations provided that $n \neq m$. As expected, relationships (150) and (151) for the propagation parameter are left unchanged for all modes having cut-off frequencies not close to the driving frequency. However, if there is a mode $m \geqslant 1$ such that Eq. (152) is fulfilled, Eqs. (161) and (162) lead to the interesting result that the propagation parameter $\hat{\Gamma}_m$ is then given by

$$\hat{\Gamma}_m = \alpha^{1/2} j(2\hat{\bar{F}}_m + 2\omega_m\omega_\alpha)^{1/2} + \alpha^{3/2}\hat{\Gamma}_{m\alpha^{3/2}} + \cdots, \tag{163}$$

which is valid for all frequencies ω arbitrarily near the cut-off frequency $a\gamma_m$, i.e., even for $\omega_\alpha = 0$. This solution agrees perfectly with the approximation derived by Hudde,[24] using the concept of boundary layer admittance; the experiments reported in this paper are also in good accordance with the theoretically predicted attenuation rate. In principle, the higher order correction term $\hat{\Gamma}_{m\alpha^{3/2}}$ could be calculated by evaluating the resonant forcing terms appearing in the equations governing the pressure perturbation $\hat{p}_{\alpha^2}$. Here, however, a different (more intuitive) approach shall be used: In contrast to the limits of the original solvabilty conditions (133) and (144) as $\hat{k}_n \to 0$, $n=1,2,...$, the corresponding limits

of the propagation parameters given by Eq. (150) can be calculated without any difficulties. As it turns out, the leading order term of $\hat{\Gamma}_m$, which is of the order $O(\alpha^{1/2})$, can also be derived by substituting for ω from Eq. (152) in Eq. (150), setting $n=m$, and performing a series expansion with respect to the perturbation parameter α. Consequently, it is very reasonable to assume that for each mode, the expression for $\hat{\Gamma}_n$ from Eq. (150) *including* the $O(\alpha^2)$ terms retains its validity even for those driving frequencies that are very close to the cut-off frequency of this mode, which enables the computation of the higher order correction term of $\hat{\Gamma}_m$. Hence, one obtains

$$\hat{\Gamma}_{m\alpha^{3/2}}=\frac{j}{2}(2\hat{\tilde{F}}_m+2\omega_m\omega_\alpha)^{1/2}\times\left(\frac{\omega_\alpha}{2\omega_m}+\frac{\dfrac{\omega_\alpha}{\omega_m}\hat{\tilde{F}}_m+\sqrt{2\omega_m}\omega_\alpha\dfrac{1-j}{d}+\hat{\tilde{K}}_m}{\hat{\tilde{F}}_m+\omega_m\omega_\alpha}\right), \tag{164}$$

where

$$\hat{\tilde{K}}_m=\lim_{\omega\to\omega_m}k_m^2\left(\hat{H}_m+\frac{\hat{F}_m^2}{2}\right)=j\omega_m\frac{\gamma-1}{2d^2}\left(\frac{1}{\mathrm{Pr}}-\frac{4}{\sqrt{\mathrm{Pr}}}\right)-\frac{j\omega_m^3}{2a^2d^2}\left(\frac{4}{3}+\eta+\frac{\gamma-1}{\mathrm{Pr}}\right). \tag{165}$$

The corresponding result obtained when the general expression for the propagation parameters from Eq. (150) is rewritten in terms of dimensional quantities can be found in Appendix C. As far as the fundamental mode is concerned it agrees perfectly with the solution given by Kergomard.[25] This applies even for the terms appearing in $\hat{H}_0$ that result from heat conduction and viscosity in the core region. Unfortunately, in that paper, no derivations were presented. Later Bruneau *et al.*[26] calculated the propagation parameters for the higher order modes, starting from a generalized dispersion equation. However, in the intermediate steps that then followed only boundary layer effects were taken into account (the mentioned dispersion relation was corrected in a subsequent paper, see Ref. 27; if, though, only the axisymmetric modes are considered, as is the case here, it remains unchanged). The expressions for the propagation parameters $\hat{\Gamma}_n$ presented in Eqs. (150) and (C4) are in complete accordance with the results in Ref. 26, if the terms proportional to $1/a^2$ appearing in the definitions (145) are formally omitted, so that the quantities $\hat{H}_n$ reduce to $\hat{H}_n=\hat{G}_n-\hat{F}_n^2/2$, $n=0,1,2,\ldots$. However, it should be kept in mind that within the framework of a correct asymptotic analysis such *a posteriori* simplifications are not appropriate, since in the case of $l=O(1)$ the terms in $\hat{H}_n$ resulting from boundary layer attenuation and those due to the viscous and thermal effects in the core region are of the same order of magnitude.

As can be seen from Eqs. (C4) and (C5), the quantity $\hat{\Gamma}_n\omega/c_0$ can be conveniently expressed in terms of the parameters St, He, Pr, γ, and η_0/μ_0. The real part of this function evaluated for the fundamental mode is displayed in Fig. 4, which clearly shows that the attenuation rate is significantly affected by the viscothermal damping mechanisms in the core region as soon as the Helmholtz number becomes an $O(1)$ quantity. In addition, Fig. 5 displays the graphs of $\Re(\hat{\Gamma}_n\omega/c_0)$ for $n=0$, 1, 2, 3 and $R=0.001$ m as functions of the Helmholtz number $\mathrm{He}=O(1)$. Here, the (large) Stokes number has been eliminated using the relationship $\mathrm{St}=\sqrt{\mathrm{He}\,\mathrm{Re}_R}$, where $\mathrm{Re}_R=c_0R\rho_0/\mu_0\sim\mathrm{Re}\gg1$ denotes the radial Reynolds number. The results plotted in this figure well illustrate the behavior of the modal damping rates when the frequency is increased such that a new mode becomes propagational: In conformity with Eqs. (151) and (163), $\Re(\hat{\Gamma}_n\omega/c_0)=O(1)$ if $\mathrm{He}<\gamma_n$ and $\gamma_n-\mathrm{He}=O(1)$, $\Re(\hat{\Gamma}_n\omega/c_0)=O(\mathrm{St}^{-1/2})$ if $|\mathrm{He}-\gamma_n|\ll1$, and $\Re(\hat{\Gamma}_n\omega/c_0)=O(\mathrm{St}^{-1})$ if $\mathrm{He}>\gamma_n$ and $\mathrm{He}-\gamma_n=O(1)$.

A further point of interest is the calculation of the transfer functions that relate the pressure fluctuations at both ends

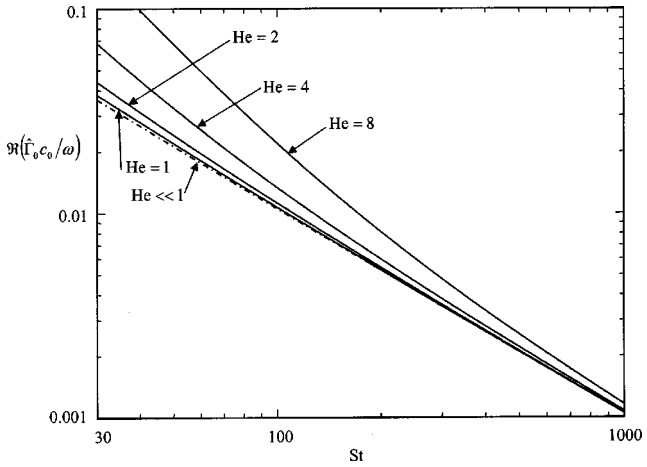


FIG. 4. Graphs of $\Re(\hat{\Gamma}_0c_0/\omega)$ as functions of St with He as parameter in double logarithmic scale; Pr=0.707, γ=1.402, η_0/μ_0=0.6 (air at 300 K); the curve for $\mathrm{He}\sim l\ll1$ corresponds to the long wavelength solution from Sec. IV; $\mathrm{St}\gg1$: $\Re(\hat{\Gamma}_0c_0/\omega)=O(\mathrm{St}^{-1})$.

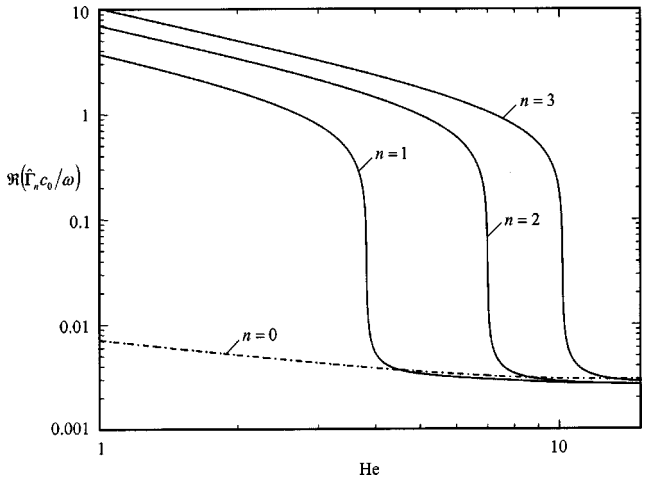


FIG. 5. Graphs of $\Re(\hat{\Gamma}_nc_0/\omega)$ as functions of He in double logarithmic scale, $R=0.001$ m, i.e., $\mathrm{St}=\sqrt{\mathrm{He}\,\mathrm{Re}_R}$ with $\mathrm{Re}_R=c_0R\rho_0/\mu_0$; Pr=0.707, γ=1.402, $\eta_0/\mu_0=0.6$, $c_0\rho_0/\mu_0=221.3732\times10^5\ \mathrm{m}^{-1}$ (air at 300 K); $\gamma_1=3.8317$, $\gamma_2=7.0156$, $\gamma_3=10.1735$.

of a tube to the axial velocities (in the case of $l \ll 1$ these relations can easily be derived by solving the system of equations given by the transfer matrix $\hat{\mathbf{A}}$ for the sound pressures $\hat{p}_{s-}$ and $\hat{p}_{s+}$). Again, the tube is regarded as a transmission line whose length is large compared to the characteristic wavelength such that $L = L_1/\alpha$ with $L_1 = O(1)$. Due to the viscothermal processes taking place in the boundary layer, each transfer function associated with a distinct mode will turn out to be affected by all the other modes as well.

For convenience, the axial velocities in the core region at both ends of the duct are assumed to involve only terms of order $O(1)$. Decomposition through the radial eigenfunctions $\mathrm{J}_0(\gamma_n r)$ then results in

$$\begin{aligned}\frac{\hat{v}_{z-}}{M} &= \hat{v}_{z1-} = \sum_{n=0}^{\infty} \hat{C}_{nv_z-} \frac{J_0(\gamma_n r)}{J_0(\gamma_n)},\\ \frac{\hat{v}_{z+}}{M} &= \hat{v}_{z1+} = \sum_{n=0}^{\infty} \hat{C}_{nv_z+} \frac{J_0(\gamma_n r)}{J_0(\gamma_n)}.\end{aligned} \tag{166}$$

Setting the origin of the axial coordinate z to the left end of the tube, evaluation of Eqs. (114) and (124) in combination with the solutions (134), (147), and (148) of the solvability conditions leads to

$$\begin{aligned}\hat{C}_{1n1} - \hat{C}_{2n1} &= \hat{C}_{nv_z-} \frac{\omega}{\hat{k}_n},\\ \hat{C}_{1n1} + \hat{C}_{2n1} &= \left(\hat{C}_{nv_z-} \coth(\hat{\Gamma}_n L) - \frac{\hat{C}_{nv_z+}}{\sinh(\hat{\Gamma}_n L)} \right) \frac{\omega}{\hat{k}_n},\end{aligned} \tag{167}$$

where, for an evanescent mode or a mode having a cut-off frequency close to the driving frequency, the terms $\coth(\hat{\Gamma}_n L)$ and $1/\sinh(\hat{\Gamma}_n L)$ can simply be replaced with 1 and 0, respectively, since in such cases the resulting corrections become exponentially small. Furthermore, annihilating the second- order solutions $\hat{v}_{z\alpha-}$ and $\hat{v}_{z\alpha+}$ given by Eqs. (115) and (130) yields

$$\begin{aligned}\hat{C}_{10\alpha} - \hat{C}_{20\alpha} &= -\hat{C}_{0v_z-} \left(\hat{F}_0 + \frac{\omega^2 \hat{F}_0}{4a^2} \right) + \sum_{m=1}^{\infty} \hat{C}_{mv_z-} \frac{2\hat{k}_m^2 \hat{F}_m}{a^2 \gamma_m^2},\\ \hat{C}_{10\alpha} + \hat{C}_{20\alpha} &= (\hat{C}_{10\alpha} - \hat{C}_{20\alpha}) \coth(\hat{\Gamma}_0 L) + \left[\hat{C}_{0v_z+} \left(\hat{F}_0 + \frac{\omega^2 \hat{F}_0}{4a^2} \right) - \sum_{m=1}^{\infty} \hat{C}_{mv_z+} \frac{2\hat{k}_m^2 \hat{F}_m}{a^2 \gamma_m^2} \right] \frac{1}{\sinh(\hat{\Gamma}_0 L)},\\ \hat{C}_{1n\alpha} - \hat{C}_{2n\alpha} &= -\hat{C}_{nv_z-} \frac{\omega \hat{F}_n}{\hat{k}_n} - \sum_{\substack{m=0\\ m\neq n}}^{\infty} \hat{C}_{mv_z-} \frac{2\omega \hat{k}_m^2 \hat{F}_m}{a^2 \hat{k}_n (\gamma_n^2 - \gamma_m^2)},\end{aligned} \tag{168}$$

$$\begin{aligned}\hat{C}_{1n\alpha} + \hat{C}_{2n\alpha} &= (\hat{C}_{1n\alpha} - \hat{C}_{2n\alpha}) \coth(\hat{\Gamma}_n L) + \left(\hat{C}_{nv_z+} \frac{\omega \hat{F}_n}{\hat{k}_n} + \sum_{\substack{m=0\\ m\neq n}}^{\infty} \hat{C}_{mv_z+} \frac{2\omega \hat{k}_m^2 \hat{F}_m}{a^2 \hat{k}_n (\gamma_n^2 - \gamma_m^2)} \right) \frac{1}{\sinh(\hat{\Gamma}_n L)},\\ & n = 1,2,\ldots,\end{aligned}$$

where the relationship

$$\begin{aligned}&\int_0^1 2r \frac{\mathrm{J}_0(\gamma_n r)}{\mathrm{J}_0(\gamma_n)} \frac{r\mathrm{J}_1(\gamma_m r)}{\gamma_m \mathrm{J}_0(\gamma_m)} dr\\ &= \begin{cases} \dfrac{2}{\gamma_n^2 - \gamma_m^2}, & n = 0,1,2\ldots, \quad m = 1,2\ldots, \quad m \neq n\\ 0, & m = n = 1,2\ldots \end{cases}\end{aligned} \tag{169}$$

has been used.

Equations (124), (130), (167), and (168) together with Eqs. (134), (147), and (148) are sufficient to determine the solutions

$$\begin{aligned}\frac{\hat{p}_{s-}}{M} &= \hat{p}_{1-} + \alpha \hat{p}_{\alpha-} + \cdots = \sum_{n=0}^{\infty} \hat{C}_{np_s-} \frac{J_0(\gamma_n r)}{J_0(\gamma_n)},\\ \frac{\hat{p}_{s+}}{M} &= \hat{p}_{1+} + \alpha \hat{p}_{\alpha+} + \cdots = \sum_{n=0}^{\infty} \hat{C}_{np_s+} \frac{J_0(\gamma_n r)}{J_0(\gamma_n)}\end{aligned} \tag{170}$$

at $z = 0$ and $z = L$, respectively, in terms of the radial eigenfunctions. The thus obtained expressions for the coefficients of the eigenfunction expansions can be written in the form

$$\begin{aligned}\hat{C}_{np_s\mp} &= \left(\pm \hat{C}_{nv_z\mp} \coth(\hat{\Gamma}_n L) \mp \frac{\hat{C}_{nv_z\pm}}{\sinh(\hat{\Gamma}_n L)} \right) \times (1 - \alpha \hat{F}_n) \frac{\omega}{\hat{k}_n} \mp \alpha \sum_{\substack{m=0\\ m\neq n}}^{\infty} \hat{C}_{mv_z\mp} \frac{2\hat{k}_m^2 \hat{F}_m}{a^2(\gamma_n^2 - \gamma_m^2)}\\ &\quad \times \left(\coth(\hat{\Gamma}_n L) \frac{\omega}{\hat{k}_n} - \coth(\hat{\Gamma}_m L) \frac{\omega}{\hat{k}_m} \right)\\ &\quad \pm \alpha \sum_{\substack{m=0\\ m\neq n}}^{\infty} \hat{C}_{mv_z\pm} \frac{2\hat{k}_m^2 \hat{F}_m}{a^2(\gamma_n^2 - \gamma_m^2)}\\ &\quad \times \left(\frac{1}{\sinh(\hat{\Gamma}_n L)} \frac{\omega}{\hat{k}_n} - \frac{1}{\sinh(\hat{\Gamma}_m L)} \frac{\omega}{\hat{k}_m} \right) + \cdots,\\ &\quad n = 0,1,2,\ldots .\end{aligned} \tag{171}$$

As pointed out earlier, the terms $\coth(\cdot)$ and $1/\sinh(\cdot)$ appearing in Eq. (171) can simply be replaced with 1 and 0, respectively, if n or m correspond to an evanescent mode or a mode having a cut-off frequency close to the driving frequency, which reveals the well-known fact that the pressure perturbations at one end can only be affected by velocity fluctuations at the other end that are associated with the propagating modes n, $m = 0,1,2,\ldots,q$, where q denotes the highest mode such that $\hat{k}_q$ is a positive real of order $O(1)$. Therefore, as far as the propagating modes are concerned, the pressure

fluctuations at one end can also be expressed in terms of the components of the axial velocity and the sound pressure at the other side by rearranging the systems (171), substituting the resulting equations mutually, and truncating after the second-order terms. It then follows that

$$\hat{C}_{np_{s\mp}} = \hat{C}_{np_{s\pm}} \cosh(\hat{\Gamma}_n L) \pm \hat{C}_{nv_{z\pm}} \sinh(\hat{\Gamma}_n L)(1-\alpha\hat{F}_n)\frac{\omega}{\hat{k}_n} + \alpha \sum_{\substack{m=0 \\ m\neq n}}^{q} \hat{C}_{mp_{s\pm}} \frac{2\hat{k}_m^2 \hat{F}_m}{a^2(\gamma_n^2-\gamma_m^2)} (\cosh(\hat{\Gamma}_n L) - \cosh(\hat{\Gamma}_m L)) \mp \alpha \sum_{\substack{m=0 \\ m\neq n}}^{q} \hat{C}_{mv_{z\pm}} \frac{2\hat{k}_m^2 \hat{F}_m}{a^2(\gamma_n^2-\gamma_m^2)} \times \left(\sinh(\hat{\Gamma}_n L)\frac{\omega}{\hat{k}_n} - \sinh(\hat{\Gamma}_m L)\frac{\omega}{\hat{k}_m} \right) + \cdots, \quad n = 0,1,2,\ldots,q. \tag{172}$$

The expressions from Eqs. (171) and (172) written in dimensional form are presented in Appendix C.

In order to derive the asymptotically correct expressions for the volume flows at both ends, the axial Navier–Stokes equation (17) is averaged over the cross section and expanded with respect to the perturbation parameter α, yielding

$$j\omega\hat{\bar{v}}_{z1} + \frac{\partial \hat{\bar{p}}_1}{\partial z} = 0, \quad j\omega\hat{\bar{v}}_{z\alpha} + \frac{2}{d^2}\frac{\partial \hat{V}_{z1}}{\partial s}\bigg|_{s=0} + \frac{\partial \hat{\bar{p}}_\alpha}{\partial z} + \frac{\partial \hat{\bar{p}}_1}{\partial z_1} = 0. \tag{173}$$

Since the expansion terms of the pressure fluctuations in the boundary layer $\hat{P}_1$ and $\hat{P}_\alpha$ are independent of the inner coordinate s, the quantities $\hat{\bar{p}}_1$ and $\hat{\bar{p}}_\alpha$ can be conveniently calculated by integrating the inner solutions from Eqs. (124) and (130) over the core region. Hence, using Eqs. (127), (167), and (168) results in

$$\frac{\hat{u}_\mp}{M} = \frac{\hat{\bar{v}}_{z\mp}}{M} = \hat{\bar{v}}_{z1\mp} + \alpha\hat{\bar{v}}_{z\alpha\mp} + \cdots = \hat{C}_{0v_{z\mp}} - \alpha\frac{1-j}{d}\sqrt{\frac{2}{\omega}}\sum_{n=0}^{\infty}\hat{C}_{nv_{z\mp}} + \cdots. \tag{174}$$

In the limit as $a\to\infty$, i.e., $l\ll 1$, the solutions for the fundamental mode from Eq. (171) then reduce to

$$\hat{\bar{p}}_{s\mp} = M\hat{C}_{0p_{s\mp}} = \left(\pm\hat{u}_\mp \coth(\hat{\Gamma}_0 L) \mp \frac{\hat{u}_\pm}{\sinh(\hat{\Gamma}_0 L)} \right) \times \left[1 + \alpha\frac{1-j}{d\sqrt{2\omega}} \left(1 - \frac{\gamma-1}{\sqrt{\mathrm{Pr}}} \right) \right] + \cdots \tag{175}$$

and the matrix system (4), with the parameters $\hat{Z}_c$ and $\hat{\Gamma}$ given by Eq. (77), is recovered.

VI. CONCLUSIONS

In the work presented here the method of multiple scales in combination with a matched asymptotic analysis has been carried out to provide insight into the linear evolution of sound pressure waves in long hard-walled ducts. The calculations primarily proceed from the assumptions that the fluid is a perfect gas and, additionally, that the acoustic boundary layer is thin compared to the characteristic dimension of the cross-sectional area. Furthermore, two different assumptions concerning the diameter to wavelength ratio have been adopted in order to derive the transfer characteristics of a long tube up to the second-order terms.

In the first case, where the wavelengths are assumed to be large compared to the diameter, it is found that the small perturbation parameter ϵ introduced by that ratio plays only a passive role, that is to say, its smallness prevents higher order modes from being excited. As a consequence, the series expansion of the characteristic impedance up to second order and the series expansion of the propagation parameter, which has to be calculated up to the third order, depend only on the (small) scaling parameter α determining the thickness of the boundary layer in terms of the diameter. However, as shown in Secs. IV B and IV C, the parameter ϵ will become important if the effects resulting from radiation at the tube end are to be incorporated into the analysis.

In Sec. V the reduced frequency (or, equivalently, the Helmholtz number) is presupposed to be of order $O(1)$, i.e., $\epsilon=1$. As a consequence, the analysis has to account for the excitation of higher order modes and the interaction of the different modes in the acoustic boundary layer. To demonstrate this, the transfer functions relating the sound pressures at both ends of the duct to the axial velocities have been calculated. The series expansions derived for the propagation parameters extend the results given in the literature with additional terms resulting from shear and bulk viscosity and heat conduction in the core region. In addition, special emphasis has been placed on the asymptotically correct treatment of modes having cut-off frequencies close to the driving frequency.

Obviously, some of the simplifying assumptions made here can be relaxed in order to account for the physical mechanisms neglected so far. Examples include nonlinear effects and the consideration of the asymmetric modes.

APPENDIX A: DUCTS WITH RECTANGULAR CROSS SECTIONS

The derivations of $\hat{Z}$ and $\hat{Y}$ for the case of a rectangular tube, which were elaborated by Stinson[8] (see also Roh *et al.*[28]), result from his general procedure developed for ducts having arbitrary cross-sectional shape. As explicated in Ref. 8, this theory is primarily based on the hypotheses that both the characteristic wavelength λ as well as the inverse of the propagation parameter $\hat{\Gamma}=\sqrt{\hat{Z}\hat{Y}}$ are very large compared to the boundary layer thickness δ, the density perturbations and the sound pressure are of comparable magnitude when scaled by their equilibrium values, and, furthermore, the sound pressure does not vary significantly through the cross section. This set of assumptions then enables the simplification of the basic equations such that, in addition to the case of circular tubes, even in the case of tubes with rectangular

cross section, the expressions for the shunt admittance and the series impedance can be given in closed form:

Let $2h$ and $2b$ be the height and the width of the tube with $D=2h$ as the characteristic dimension of the cross section and $b/h=O(1)$. Additionally, the following parameters

$$a_k=\frac{(k+1/2)\pi}{h},\qquad b_n=\frac{(n+1/2)\pi}{b} \tag{A1}$$

are introduced. Then the quantities $\hat{Z}$ and $\hat{Y}$ are constructed from

$$\hat{Z}=\frac{\mu_0 b^2 h^2}{4S\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{1}{a_k^2 b_n^2\left(a_k^2+b_n^2+\frac{j\omega\rho_0}{\mu_0}\right)}}, \tag{A2}$$

$$\hat{Y}=\frac{j\omega S}{p_0}\left[1-\frac{4(\gamma-1)j\omega\rho_0 C_p}{\gamma\kappa_0 b^2 h^2}\times\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{1}{a_k^2 b_n^2\left(a_k^2+b_n^2+\frac{j\omega\rho_0 C_p}{\kappa_0}\right)}\right],$$

where $S=4bh$.

Since for every arbitrary $\hat{\xi}$ the relationship

$$\sum_{k=0}^{\infty}\frac{1}{h^2 a_k^2(h^2 a_k^2+\hat{\xi})}=\frac{1}{2\hat{\xi}}\left(1-\frac{\tanh(\sqrt{\hat{\xi}})}{\sqrt{\hat{\xi}}}\right) \tag{A3}$$

holds, expression (A2) can be recast into

$$\hat{Z}=\frac{\mu_0 b^2}{2Sh^2\sum_{n=0}^{\infty}\frac{1}{b_n^2\hat{\xi}_1}\left(1-\frac{\tanh(\sqrt{\hat{\xi}_1})}{\sqrt{\hat{\xi}_1}}\right)}, \tag{A4}$$

$$\hat{Y}=\frac{j\omega S}{p_0}\left[1-\frac{2(\gamma-1)j\omega\rho_0 C_p h^2}{\gamma\kappa_0 b^2}\times\sum_{n=0}^{\infty}\frac{1}{b_n^2\hat{\xi}_2}\left(1-\frac{\tanh(\sqrt{\hat{\xi}_2})}{\sqrt{\hat{\xi}_2}}\right)\right].$$

Here, $\hat{\xi}_1$ and $\hat{\xi}_2$ are abbreviations for

$$\hat{\xi}_1=h^2\left(b_n^2+\frac{j\omega\rho_0}{\mu_0}\right),\qquad \hat{\xi}_2=h^2\left(b_n^2+\frac{j\omega\rho_0 C_p}{\kappa_0}\right). \tag{A5}$$

Due to the infinite sums in Eqs. (A2) and (A4) the expansions in the limits as $\mathrm{St}\to 0$ (LFL) or $\mathrm{St}\to\infty$ (HFL), with the Stokes number defined as

$$\mathrm{St}=\sqrt{\frac{\omega\rho_0 h^2}{\mu_0}}, \tag{A6}$$

turn out to be very tedious. However, one could try to derive at least the asymptotically correct expressions for Eq. (A4) in the HFL by means of a multiple scales analysis based on the *low reduced frequency assumptions* stated in Eq. (5) in combination with a matched asymptotic expansion, as outlined in Sec. IV. Such an approach involving separate calculations for the acoustic flow in the core region, the main boundary regions, and, in principle, also in the corner regions of the boundary layer has the advantage that the series expansions of infinite sums can then be avoided. In contrast to the study in Sec. IV, the following analysis confines itself to investigating the *leading* order terms of the volume flow $\hat{u}$ and the sound pressure $\hat{p}_s$ generated by waves propagating over a distance $L\sim\mathrm{St}\,\lambda$ or, equivalently, the *leading* order terms and the *second*-order terms of the quantities $\hat{u}$ and $\hat{p}_s$ generated by waves propagating over a spatial range $L\sim\lambda$ only.

To this end, the nondimensional coordinates

$$x^*=\frac{x}{b},\quad y^*=\frac{y}{h} \tag{A7}$$

for the vertical and the horizontal direction, respectively, and the small scaling parameters

$$\epsilon=a\frac{h}{\lambda}\sim l,\quad \alpha=d\frac{\lambda}{h\sqrt{\mathrm{Re}}}\sim\mathrm{St}^{-1} \tag{A8}$$

are introduced. Omitting the superscripts $*$ denoting nondimensional quantities, the boundary layer coordinates are given by

$$s_{\pm y}=\frac{1\mp y}{\alpha},\quad s_{\pm x}=\frac{1\mp x}{\alpha}. \tag{A9}$$

As mentioned earlier, in this simplified analysis either the changes of the first- and second-order terms of the quantities $\hat{p}_s$ and $\hat{u}$ are to be resolved over a spatial range of the order $O(1)$ or just the leading order effects emerging over distances of the order $O(\alpha^{-1})$ are to be investigated. In both cases two length scales are sufficient, resulting in

$$\frac{\partial}{\partial z} \rightarrow \frac{\partial}{\partial z} + \alpha \frac{\partial}{\partial z_1}. \tag{A10}$$

In order to guarantee the linearity of the problem, the Mach number now has to be assumed to satisfy

$$M \ll \epsilon, \quad M \ll \alpha. \tag{A11}$$

Furthermore, it should be noted that the effects resulting from the viscothermal processes taking place in the very small corner regions of the boundary layer where both $s_{\pm x}$ as well as $s_{\pm y}$ are of order $O(1)$ will not enter the correction terms at second order in α, since their contribution to the volume flow is an order smaller than that of the main part of the boundary layer. Hence, deriving an exact solution of the boundary layer equations for the corner regions is not necessary.

A perturbation analysis for the HFL very similar to that carried out in Sec. IV then shows that the series impedance and the shunt admittance in dimensional form assume the limiting values

$$\hat{Z} = \frac{\omega \rho_0}{S}\left[j + (1+j)\left(1 + \frac{h}{b}\right)\sqrt{\frac{\mu_0}{2\omega\rho_0 h^2}} + O(\mathrm{St}^{-2})\right], \tag{A12}$$

$$\hat{Y} = \frac{\omega S}{\gamma p_0}\left[j + (1+j)\left(1 + \frac{h}{b}\right)(\gamma - 1)\sqrt{\frac{\mu_0}{2\omega\rho_0 h^2}\frac{\kappa_0}{\mu_0 C_p}} + O(\mathrm{St}^{-2})\right].$$

The low reduced frequency assumptions from Eq. (5) together with $\mathrm{St} \sim h/\delta \gg 1$ yield the ordering relationship $\lambda \gg h \gg \delta$ and Eq. (A12) implies that in the HFL, the inverse of the dimensional propagation parameter is of the order $O(\lambda)$. Furthermore, as in the case of tubes with circular cross-sectional shape, the sound pressure does not change significantly over the cross section. As a consequence, the basic assumptions adopted by Stinson (see above) are satisfied, which leads to the conclusion that Eq. (A12) represents the asymptotically correct approximations to Eq. (A2) in the limit of large Stokes numbers.

APPENDIX B: DUCTS WITH SLIT-SHAPED AND ARBITRARY CROSS SECTIONS

In the case of a rectangular slit-shaped tube with $D = 2h$ as the characteristic dimension such that $h/b \ll 1$, the expressions for $\hat{Z}$ and $\hat{Y}$ derived for $b/h = O(1)$ reduce to[8]

$$\hat{Z} = \frac{j\omega\rho_0}{S\left(1 - \sqrt{\frac{\mu_0}{j\omega\rho_0}}\frac{\tanh\left(h\sqrt{\frac{j\omega\rho_0}{\mu_0}}\right)}{h}\right)}, \tag{B1}$$

$$\hat{Y} = \frac{j\omega S\left[1 + (\gamma - 1)\sqrt{\frac{\kappa_0}{j\omega\rho_0 C_p}}\frac{\tanh\left(h\sqrt{\frac{j\omega\rho_0 C_p}{\kappa_0}}\right)}{h}\right]}{\gamma p_0},$$

with $S = 4bh$. The result for $\hat{Z}$ is in accordance with the solution already given in the 1975 paper by Backus;[29] however, please note that in the expression for the shunt admittance stated there, a factor of $\sqrt{j}$ is missing. Derivations of $\hat{Z}$ and $\hat{Y}$ can also be found in Ingard[30] (pp. 2/19–2/29).

The LFL can be calculated by means of a power series expansion with respect to the Stokes number from Eq. (A6) and reads

$$\begin{aligned} &\text{LFL:} \quad \hat{Z} = \frac{3\mu_0}{h^2 S}\left(1 + j\frac{2\omega\rho_0 h^2}{5\mu_0} + O(\mathrm{St}^4)\right), \\ &\hat{Y} = \frac{\omega S}{p_0}\left(j + \frac{\gamma - 1}{\gamma}\frac{\omega\rho_0 h^2}{3\mu_0}\frac{\mu_0 C_p}{\kappa_0} + O(\mathrm{St}^4)\right). \end{aligned} \tag{B2}$$

Furthermore, since for $\mathrm{St} \rightarrow \infty$

$$\tanh(\hat{\xi}) \rightarrow 1, \tag{B3}$$

expression (B1) can be evaluated to give the HFL in the form

$$\begin{aligned} &\text{HFL:} \quad \hat{Z} = \frac{\omega\rho_0}{S}\left[j + (1+j)\sqrt{\frac{\mu_0}{2\omega\rho_0 h^2}} + O(\mathrm{St}^{-2})\right], \\ &\hat{Y} = \frac{\omega S}{\gamma p_0}\left[j + (1+j)(\gamma - 1)\sqrt{\frac{\mu_0}{2\omega\rho_0 h^2}\frac{\kappa_0}{\mu_0 C_p}} + O(\mathrm{St}^{-2})\right]. \end{aligned} \tag{B4}$$

As expected, the relationships (A12) as well as Eqs. (B4) and (14) confirm the assumption stated in Morse and Ingard[14] (p. 475) and elsewhere that for $\mathrm{St} \rightarrow \infty$ terms resulting from viscous dissipation ($\hat{Z}$) and heat conduction ($\hat{Y}$) in the boundary layer are always proportional to P/S, with P being the perimeter of the cross-sectional area (however, in contrast to the results for $\hat{Z}$ and $\hat{Y}$ presented here, in Ref. 14 the effects of heat conduction are contained within the resistive part of the series impedance). For an arbitrary cross section the HFL can thus be written as

$$\begin{aligned} &\text{HFL:} \quad \hat{Z} = \frac{\omega\rho_0}{S}\left[j + (1+j)\frac{P}{S}\sqrt{\frac{\mu_0}{2\omega\rho_0}} + \cdots\right], \\ &\hat{Y} = \frac{\omega S}{\gamma p_0}\left[j + (1+j)\frac{P}{S}(\gamma - 1)\sqrt{\frac{\mu_0}{2\omega\rho_0}\frac{\kappa_0}{\mu_0 C_p}} + \cdots\right]. \end{aligned} \tag{B5}$$

Further (semianalytical) results for the LFL and HFL in the case of rectangular and other cross sections, which were derived by using variational methods, are presented in Cummings.[31]

APPENDIX C: RESULTS IN DIMENSIONAL FORM

Expressions used in the following:

$$S=R^2\pi,\quad \mathrm{St}(\omega)=\sqrt{\frac{\omega\rho_0R^2}{\mu_0}}\gg 1,$$

$$\mathrm{Pr}=\frac{\mu_0C_p}{\kappa_0}=O(1),\qquad \mathrm{He}(\omega)=\frac{\omega R}{c_0}. \tag{C1}$$

Equation (76):

$$\hat{Z}=\frac{\omega\rho_0}{S}\left[j+(1+j)\frac{\sqrt{2}}{\mathrm{St}(\omega)}+\frac{3}{\mathrm{St}(\omega)^2}+O(\mathrm{St}^{-3})\right],$$

$$\hat{Y}=\frac{\omega S}{\gamma p_0}\left[j+(1+j)\frac{(\gamma-1)}{\mathrm{St}(\omega)}\sqrt{\frac{2}{\mathrm{Pr}}}-\frac{(\gamma-1)}{\mathrm{St}(\omega)^2\,\mathrm{Pr}}+O(\mathrm{St}^{-3})\right],$$

$$\hat{Z}_c=\frac{c_0\rho_0}{S}\left[1+\frac{1-j}{\sqrt{2}\,\mathrm{St}(\omega)}\left(1-\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)-\frac{j}{2\,\mathrm{St}(\omega)^2}\times\left(2-2\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}-\frac{5\gamma-3\gamma^2-2}{\mathrm{Pr}}\right)+O(\mathrm{St}^{-3})\right], \tag{C2}$$

$$\hat{\Gamma}=\frac{j\omega}{c_0}\left[1+\frac{1-j}{\sqrt{2}\,\mathrm{St}(\omega)}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)-\frac{j}{\mathrm{St}(\omega)^2}\times\left[1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left(1-\frac{\gamma}{2\sqrt{\mathrm{Pr}}}\right)\right]+O(\mathrm{St}^{-3})\right].$$

Equation (81):

$$\hat{Z}_t=\frac{c_0\rho_0}{S}\left[\tanh\left(\frac{\omega L}{c_0}\left[j+\frac{1+j}{\sqrt{2}\,\mathrm{St}(\omega)}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\right]\right)+\frac{1-j}{\sqrt{2}\,\mathrm{St}(\omega)}\left(1-\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\tanh\left(\frac{\omega L}{c_0}\left[j+\frac{1+j}{\sqrt{2}\,\mathrm{St}(\omega)}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\right]\right)+\frac{0.8217j\,\mathrm{He}(\omega)+\frac{\omega L}{c_0\,\mathrm{St}(\omega)^2}\left[1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left(1-\frac{\gamma}{2\sqrt{\mathrm{Pr}}}\right)\right]}{\cosh\left(\frac{\omega L}{c_0}\left[j+\frac{1+j}{\sqrt{2}\,\mathrm{St}(\omega)}\left(1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right)\right]\right)^2}+O(\mathrm{St}^{-2})\right], \tag{C3}$$

where $\omega L/c_0=O(\mathrm{St})$ and $\mathrm{He}=O(\mathrm{St}^{-1})$.

Equation (150):

$$\hat{\Gamma}_n=\sqrt{-\hat{k}_n^2\left(1+2\frac{\hat{F}_n}{\mathrm{St}(\omega)}+\frac{2\hat{H}_n+\hat{F}_n^2}{\mathrm{St}(\omega)^2}\right)+O(\mathrm{St}^{-3})},\qquad n=0,1,2,..., \tag{C4}$$

where

$$\hat{k}_n=\begin{cases}1/R\sqrt{\mathrm{He}(\omega)^2-\gamma_n^2}, & \mathrm{He}(\omega)>\gamma_n\\ -j/R\sqrt{\gamma_n^2-\mathrm{He}(\omega)^2}, & \mathrm{He}(\omega)<\gamma_n\end{cases},$$

$$\hat{F}_n=\frac{1-j}{\sqrt{2}}\left(1+\frac{\mathrm{He}(\omega)^2}{\hat{k}_n^2R^2}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\right),\quad n=0,1,2,... \tag{C5}$$

$$\hat{H}_0=-j\left[1+\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left(1-\frac{\gamma}{2\sqrt{\mathrm{Pr}}}\right)\right]-\frac{j\,\mathrm{He}(\omega)^2}{2}\left[\frac{11}{6}+\frac{\eta_0}{\mu_0}+\frac{\gamma-1}{2\sqrt{\mathrm{Pr}}}\left(2+\frac{1+\gamma}{\sqrt{\mathrm{Pr}}}\right)\right],$$

$$\hat{H}_n=-j\left[1+\frac{\mathrm{He}(\omega)^2}{\hat{k}_n^2R^2}\frac{\gamma-1}{\sqrt{\mathrm{Pr}}}\left[1-\frac{1}{2\sqrt{\mathrm{Pr}}}\left(1+\frac{\mathrm{He}(\omega)^2}{\hat{k}_n^2R^2}\times(\gamma-1)\right)\right]\right]-\frac{j\,\mathrm{He}(\omega)^4}{2\hat{k}_n^2R^2}\left(\frac{4}{3}+\frac{\eta_0}{\mu_0}+\frac{\gamma-1}{\mathrm{Pr}}\right),\quad n=1,2,...,$$

and, furthermore, $\mathrm{He}=O(1)$.

Equations (171) and (172):

$$\hat{C}_{np_{s}^{\mp}} = \left(\pm \hat{C}_{nv_{z}^{\pm}} \coth(\hat{\Gamma}_n L) \mp \frac{\hat{C}_{nv_{z}^{\pm}}}{\sinh(\hat{\Gamma}_n L)} \right) \times \left(1 - \frac{\hat{F}_n}{\mathrm{St}(\omega)} \right) \frac{\mathrm{He}(\omega)}{\hat{k}_n R} \mp \sum_{\substack{m=0 \\ m \neq n}}^{\infty} \hat{C}_{mv_{z}^{\mp}} \frac{2\hat{k}_m^2 R^2 \hat{F}_m}{\mathrm{St}(\omega)(\gamma_n^2 - \gamma_m^2)} \times \left(\coth(\hat{\Gamma}_n L) \frac{\mathrm{He}(\omega)}{\hat{k}_n R} - \coth(\hat{\Gamma}_m L) \frac{\mathrm{He}(\omega)}{\hat{k}_m R} \right) \pm \sum_{\substack{m=0 \\ m \neq n}}^{\infty} \hat{C}_{mv_{z}^{\pm}} \frac{2\hat{k}_m^2 R^2 \hat{F}_m}{\mathrm{St}(\omega)(\gamma_n^2 - \gamma_m^2)} \left(\frac{1}{\sinh(\hat{\Gamma}_n L)} \frac{\mathrm{He}(\omega)}{\hat{k}_n R} - \frac{1}{\sinh(\hat{\Gamma}_m L)} \frac{\mathrm{He}(\omega)}{\hat{k}_m R} \right) + \cdots, \quad n = 0,1,2,..., \tag{C6}$$

$$\hat{C}_{np_{s}^{\mp}} = \hat{C}_{np_{s}^{\pm}} \cosh(\hat{\Gamma}_n L) \pm \hat{C}_{nv_{z}^{\pm}} \sinh(\hat{\Gamma}_n L) \times \left(1 - \frac{\hat{F}_n}{\mathrm{St}(\omega)} \right) \frac{\mathrm{He}(\omega)}{\hat{k}_n R} + \sum_{\substack{m=0 \\ m \neq n}}^{q} \hat{C}_{mp_{s}^{\pm}} \frac{2\hat{k}_m^2 R^2 \hat{F}_m}{\mathrm{St}(\omega)(\gamma_n^2 - \gamma_m^2)} (\cosh(\hat{\Gamma}_n L) - \cosh(\hat{\Gamma}_m L)) \mp \sum_{\substack{m=0 \\ m \neq n}}^{q} \hat{C}_{mv_{z}^{\pm}} \frac{2\hat{k}_m^2 R^2 \hat{F}_m}{\mathrm{St}(\omega)(\gamma_n^2 - \gamma_m^2)} \times \left(\sinh(\hat{\Gamma}_n L) \frac{\mathrm{He}(\omega)}{\hat{k}_n R} - \sinh(\hat{\Gamma}_m L) \frac{\mathrm{He}(\omega)}{\hat{k}_m R} \right) + \cdots, \quad n = 0,1,2,...,q, \tag{C7}$$

where $\hat{\Gamma}_n$, $\hat{k}_n$, and $\hat{F}_n$ are defined as in Eqs. (C4) and (C5), respectively, $\mathrm{He} = O(1)$ and $\omega L/c_0 = O(\mathrm{St})$. Then the sound pressure, the axial velocity, the average sound pressure, and the volume flow at the ends of the tube are given by

$$\begin{aligned} \hat{p}_{s\mp} &= c_0 \rho_0 \sum_{n=0}^{\infty} \hat{C}_{np_{s}^{\mp}} \frac{\mathrm{J}_0(\gamma_n r/R)}{\mathrm{J}_0(\gamma_n)}, \\ \hat{v}_{z\mp} &= \sum_{n=0}^{\infty} \hat{C}_{nv_{z}^{\mp}} \frac{\mathrm{J}_0(\gamma_n r/R)}{\mathrm{J}_0(\gamma_n)}, \\ \hat{\bar{p}}_{s\mp} &= c_0 \rho_0 C_{0p_{s}^{\mp}}, \\ \hat{u}_{\mp} &= \hat{C}_{0v_{z}^{\mp}} - (1-j) \frac{\sqrt{2}}{\mathrm{St}(\omega)} \sum_{n=0}^{\infty} \hat{C}_{nv_{z}^{\mp}} + \cdots. \end{aligned} \tag{C8}$$

[1] G. Kirchhoff, "Über den einfluss der wärmeleitung in einem gase auf die schallbewegung (On the effect of heat conduction in a gas on the propagation of sound)," (in German) Ann. Phys. (Leipzig) **134**, 177–193 (1868).

[2] L. Rayleigh, *Theory of Sound II* (Macmillan, London, 1896).

[3] C. Zwikker and C. Kosten, *Sound Absorbing Materials* (Elsevier, Amsterdam, 1949).

[4] A. S. Iberall, "Attenuation of oscillatory pressures in instrument lines," J. Res. Natl. Bur. Stand. **45**, 85–108 (1950).

[5] F. B. Daniels, "On the propagation of sound waves in a cylindrical conduit," J. Acoust. Soc. Am. **22**, 563–564 (1950).

[6] W. Kraak, "Der dynamische strömungsstandwert (strömungswiderstand) kreisförmiger, kurzer Kanäle (The dynamic flow resistance of circular short ducts)," (in German) Hochfrequenztechnik Elektroakustik **65**, 46–49 (1956).

[7] H. Tijdeman, "On the propagation of sound waves in cylindrical tubes," J. Sound Vib. **39**, 1–33 (1975).

[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] J. Kergomard, "Measurement of acoustic impedance using a capillary: An attempt to achieve optimization," J. Acoust. Soc. Am. **79**, 1129–1140 (1986).

[10] D. H. Keefe, "Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximations for isothermal and nonisothermal boundary conditions," J. Acoust. Soc. Am. **75**, 58–62 (1984).

[11] A. H. Nayfeh, "Nonlinear propagation of a wave packet in a hard-walled circular duct," J. Acoust. Soc. Am. **57**, 803–809 (1975).

[12] D. G. Crighton, A. P. Dowling, J. E. Ffowcs Williams, M. Heckl, and F. G. Leppington, *Modern Methods in Analytical Acoustics* (Springer, Berlin, 1996).

[13] M. J. Anderson and P. G. Vaidya, "Application of the method of multiple scales to linear waveguide boundary value problems," J. Acoust. Soc. Am. **88**, 2450–2458 (1990).

[14] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, 1986).

[15] L. L. Beranek, *Acoustics* (Acoustical Society of America, 1996).

[16] S. N. Makarov and E. V. Vatrushina, "Effect of the acoustic boundary layer on a nonlinear quasiplane wave in a rigid-walled tube," J. Acoust. Soc. Am. **94**, 1076–1083 (1993).

[17] Q. Qi, R. E. Johnson, and J. G. Harris, "Boundary layer attenuation and acoustic streaming accompanying plane-wave propagation in a tube," J. Acoust. Soc. Am. **97**, 1499–1509 (1995).

[18] A. D. Pierce, *Acoustics* (Acoustical Society of America, 1989).

[19] L. Rayleigh, "On the theory of resonance," Philos. Trans. R. Soc. London **161**, 77–118 (1870).

[20] M. S. Howe, "On Rayleigh's computation of the 'end correction,' with application to the compression wave generated by a train entering a tunnel," J. Fluid Mech. **385**, 63–78 (1999).

[21] P. J. Daniell, "The coefficient of end-correction. I," Philos. Mag. **30**, 137–146 (1915).

[22] A. N. Norris and I. C. Sheng, "Acoustic radiation from a circular pipe with an infinite flange," J. Sound Vib. **135**, 85–93 (1989).

[23] R. E. Beatty, Jr., "Boundary layer attenuation of higher order modes in rectangular and circular tubes," J. Acoust. Soc. Am. **22**, 850–854 (1950).

[24] H. Hudde, "The propagation constant in lossy circular tubes near the cutoff frequencies of higher-order modes," J. Acoust. Soc. Am. **83**, 1311–1318 (1988).

[25] J. Kergomard, "Comments on 'Wall effects on sound propagation in tubes,' " J. Sound Vib. **98**, 149–155 (1985).

[26] A. M. Bruneau, M. Bruneau, Ph. Herzog, and J. Kergomard, "Boundary layer attenuation of higher order modes in waveguides," J. Sound Vib. **119**, 15–27 (1987).

[27] J. Kergomard, M. Bruneau, A. M. Bruneau, and Ph. Herzog, "On the propagation constant of higher order modes in a cylindrical waveguide," J. Sound Vib. **126**, 178–181 (1988).

[28] H. S. Roh, W. P. Arnott, J. M. Sabatier, and R. Raspet, "Measurement and calculation of acoustic propagation constants in arrays of air-filled rectangular tubes," J. Acoust. Soc. Am. **89**, 2617–2624 (1991).

[29] J. Backus, "Acoustic impedance of an annular capillary," J. Acoust. Soc. Am. **58**, 1078–1081 (1975).

[30] U. Ingard, *Notes on Sound Absorption Technology* (Noise Control Foundation, 1994).

[31] A. Cummings, "Sound propagation in narrow tubes of arbitrary cross-section," J. Sound Vib. **162**, 27–42 (1993).

Rayleigh–Ritz calculation of the resonant modes of a solid parallelepiped in a pressurizing fluid

D. E. Beck
Department of Physics and Astronomy, University of Hawaii-Manoa, Honolulu, Hawaii 96822

(Received 16 March 2003; revised 17 October 2003; accepted 3 November 2003)

Resonant ultrasound spectroscopy relies on comparisons of experimentally determined vibrational spectra to theoretically computed spectra for the extraction of the elastic moduli of the solid samples. To determine the pressure dependence of these moduli, resonant spectra are taken for samples pressurized by a surrounding gas and knowledge of the contribution of the surface loading of the sample by the gas is needed in order to extract the intrinsic pressure dependence of the moduli. To facilitate the required comparisons, a Rayleigh–Ritz variational calculation of the vibrational spectrum is formulated which includes the loading of the solid by the pressurizing fluid. This formalism is used to compute the effect of gas loading on the vibrational spectrum of an isotropic, solid parallelepiped. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639331]

PACS numbers: 43.20.Tb, 43.35.Cg [LLT] Pages: 556–566

I. INTRODUCTION

Resonant ultrasound spectroscopy (RUS) has proven its great utility in the determination of the elastic moduli of solids[1,2] and the temperature dependence of these parameters.[3] Extraction of these moduli from RUS experiment relies on the matching of the theoretically computed vibrational spectrum to the experimental spectrum obtained by the ultrasonic excitation of the solid. For a solid with edges and corners such as single crystals, there is no analytical method for imposing the required boundary conditions at the solid surface. Therefore, the theoretical analyses have relied on the Rayleigh–Ritz variational procedure to solve the Lagrangian for the vibrational spectra of the solids in a vacuum [''natural'' or ''free surface'' solutions]. Migliori *et al.*[2] have reviewed this theory and its application to the analyses of experimental data, and Maynard[4] provides a history and overview of RUS. The application of the RUS technique to the determination of the pressure dependence of the elastic moduli requires an accounting of the modification of the vibrational spectrum from the radiation damping due to the pressurizing gas. Hence, a theoretical analysis including gas loading is required in order to separate this contribution to the spectral changes from those due to the intrinsic pressure dependence of the solid moduli.

Zhang *et al.*[5] have published experimental data along with a theoretical analysis for small fused silica parallelepipeds pressurized with various gases. Their theoretical calculation employs a two-dimensional model (TDM) to compute the radiation impedance. This model combines the radiation damping on parallel faces to obtain the damping for the parallelepiped; therefore, it includes no coupling of adjacent faces. The radiation at a face is computed by considering a vibrating rectangular piston in an infinite baffle for which the nonuniform velocity amplitude normal to the piston is specified in order to approximate that of the particular vibrational mode of the solid.

Here, the Lagrangian for the solid is extended to incorporate the gas loading so that the vibrating solid and the impedance due to the pressurizing gas are treated in a single expression. This extension is described in the next section along with the Rayleigh–Ritz procedure for the variational calculation. The extension requires the matching of the interior solution to an outgoing pressure wave in the gas at the solid surface. To facilitate this, matching the Green's function for the outgoing spherical wave is employed to describe the pressure[6] in the gas resulting from the vibrating gas–solid interface. Since the Lagrangian fully describes the solid and the gas loading, the self-consistent solution of matrix equations resulting from the Rayleigh–Ritz procedure provides the vibrational spectrum of the pressurized solid.

The formulation permits the evaluation of the fluid-loading contributions to the vibrational spectrum for any fluid to solid density ratio. The coupling at the interface is proportional to this density ratio which is much less than 1 for moderate gas pressures and results in only small shifts of the vibrational frequencies from their vacuum values. Hence, a perturbative calculation permits the accurate evaluation of these small shifts, and this calculation is described in Sec. IV. The pressure at the gas–solid interface is obtained from a set of coupled, inhomogeneous, integral equations expressing the pressure in terms of the surface velocities. The details of the integrals required for the matrix evaluation of these equations and the surface matrix elements are summarized in Appendix A. In Sec. IV, a zero-order approximation (ZO) to an iterative solution for the surface matrix elements is also introduced. In this approximation, the pressure on a face of the solid is evaluated by considering the face to be a piston in an infinite baffle. This same evaluation of the pressure is employed in the TDM; however, the computation of the radiation impedance differ in the two approximations.

The computations, presented in Sec. V, assumes an ideal fluid [nonviscous] and an isotropic medium. These assumptions greatly simplify the computation but they are not restrictions imposed by the analysis. Further, they do not hinder the comparison with experiment, since the current experimental data[5] are for an isotropic solid and the viscosity of a gas is small.[7] The comparison of the results from the ZO calculation with those from the full calculation demonstrate

that this much simpler computation gives good estimates of the shift in frequency and damping of the vibrational spectrum. Both of these results are compared with the TDM results.

A separable scalar field based on sine and cosine functions is employed to construct the basis vectors that describe the displacements in the solid. The basis vectors and their symmetry grouping for the parallelepiped are set forth in Sec. III. This choice of basis vectors permits the efficient evaluation of the multidimensional integrals required for the surface matrix elements. These matrix elements can be evaluated using a fast Fourier routine[8] to perform the discrete Fourier transforms.[9] The details of the discrete Fourier transform and the procedure for eliminating the integrable singularity from some of the transforms needed for the surface matrix elements are contained in Appendix B.

Finally, in Sec. VI, there is a discussion of the calculation of the effects of gas loading on small, single-crystal samples. The Lagrangian formulation and the Rayleigh–Ritz variational calculation presented here are not limited to isotropic parallelepipeds, and the resulting matrix equations are as amenable to solution as the vacuum case.[2] The major complication is the evaluation of the vibrating pressure at the surface of the solid, and we provide our views concerning an efficient treatment of this problem.

II. FORMALISM

The Lagrangian for an elastic solid with volume, V, in a vacuum is

$$L_O=\int\int\int_V dV[\rho\dot{u}_i^*\dot{u}_i-c_{ij;kl}u_{i,j}^*u_{k,l}],$$

where u_i is the ith component of the complex displacement (u_i^* denotes its complex conjugate), $c_{ij;kl}$ is an element of the elastic tensor, and ρ is the density of the solid. We employ the usual conventions where repeated indices are summed over, the ''comma'' denotes differentiation with respect to the coordinate denoted by the following indices and the ''dot'' denotes differentiation with respect to time.

Requiring the linear variation with respect to $u_i^*\rightarrow u_i^*+\delta u_i^*$ to be zero results in the elastic wave equation

$$\rho\ddot{u}_i=c_{ij;kl}u_{k,jl},$$

and the ''natural'' or ''free surface'' boundary condition

$$n_jc_{ij;kl}u_{k,l}|_{S_{\hat{\mathbf{n}}}}=0,$$

on the bounding surface of the solid, $S_{\hat{\mathbf{n}}}$. Here, n_j is a component of the outward pointing unit vector normal to the surface.

For a solid immersed in a pressurizing fluid, the displacement and strain at the solid surface need to be related to these quantities in the fluid. In order to facilitate this a surface contribution is added to the Lagrangian

$$L_S=\frac{1}{2}\int\int_{S_{\hat{\mathbf{n}}}} dS[\gamma_i^*u_i^*u_i^*+\gamma_iu_iu_i].$$

Requiring $\delta(L_O+L_S)/\delta u_i^*=0$ again yields the wave equation; however, the boundary condition is now given by

$$\gamma_i^*u_i^*-n_jc_{ij;kl}u_{k,l}|_{S_{\hat{\mathbf{n}}}}=0,$$

[$\delta(L_O+L_S)/\delta u_i=0$ yields the complex conjugate of this equation]. The Lagrange multipliers, γ_i^*, are fixed by establishing the correct relations for the displacement and strain in the fluid.

In order to maintain a steady state (time independence), the energy being radiated by the outgoing, pressure wave in the fluid needs a source in the Lagrangian. This is provided by adding a volume, driving term to the Lagrangian[10]

$$L_F=\frac{\rho}{2\tau}\int\int\int_V dV[u_i^*\dot{u}_i-\dot{u}_i^*u_i].$$

The resulting driving force will not be the experimentally applied force which is only imposed at the corners of the solid. However, we are only interested in the dispersion of the vibrational modes and this dispersion is independent of the means of excitation for elastic deformations. Variation of $L=L_O+L_S+L_F$ with respect to u_i^* yields the equation of motion

$$\rho\ddot{u}_i=\frac{\rho}{\tau}\dot{u}_i+c_{ij;kl}u_{k,jl}.$$

Driving the solid at a angular frequency, ω, and employing the time dependence $u_i\propto e^{-i\omega t}$, the full Lagrangian is

$$L=\int\int\int_V dV[\rho\omega(\omega-i/\tau)u_i^*u_i-c_{ij;kl}u_{i,j}^*u_{k,l}]+\frac{1}{2}\int\int_{S_{\hat{\mathbf{n}}}} dS[\gamma_i^*u_i^*u_i^*+\gamma_iu_iu_i].$$

For isotropic solids

$$c_{ij;kl}=\lambda\delta_{ij}\delta_{kl}+\mu(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}),$$

where the mechanical properties of the medium are parametrized by the Lamé constants λ and μ.[11] Employing vector notation, the Lagrangian is

$$L=\int\int\int_V dV[\rho\omega(\omega-i/\tau)\mathbf{u}^*\cdot\mathbf{u}-\lambda(\nabla\cdot\mathbf{u}^*)(\nabla\cdot\mathbf{u})-2\mu\mathbf{e}^*:\mathbf{e}]+\frac{1}{2}\int\int_{S_{\hat{\mathbf{n}}}} dS[\mathbf{u}_\gamma^*\cdot\mathbf{u}^*+\mathbf{u}_\gamma\cdot\mathbf{u}]. \quad (1)$$

Here, $\mathbf{e}$ is the symmetric strain tensor with elements $e_{ij}=(u_{i,j}+u_{j,i})/2$ and its contraction is $\mathbf{e}^*:\mathbf{e}=e_{ji}e_{ji}$. Our solid is immersed in a nonviscous fluid, density ρ_o, and Lamé constants $\mu_o=0$ (Ref. 12) and λ_o. If we place the bounding surface in the fluid, the boundary condition is

$$\mathbf{u}_\gamma^*-\hat{\mathbf{n}}\lambda_o\nabla\cdot\mathbf{u}^o|_{S_{\hat{\mathbf{n}}}}=0, \quad (2)$$

where $\mathbf{u}_\gamma^*$ has components $\gamma_i^*u_i^*$ and $\mathbf{u}^o$ is the outgoing displacement vector in the fluid. Additionally, the normal com-

ponent of the displacement must be continuous across the solid–fluid interface.

The Rayleigh–Ritz variational procedure for the minimization of L is implemented by selecting a set of basis vectors, $\mathbf{v}_\alpha(\mathbf{r})$, setting

$$\mathbf{u}(\mathbf{r})=\sum_\alpha A_\alpha \mathbf{v}_\alpha(\mathbf{r}),$$

and requiring that the linear variation of L with respect to $A_\beta^* \rightarrow A_\beta^* + \delta A_\beta^*$ vanish. This variation yields

$$\delta L/\delta A_\beta^* = \rho \sum_\alpha [\omega(\omega - i/\tau) M_{\beta\alpha} - N_{\beta\alpha} + S_{\beta\alpha}] A_\alpha = 0. \tag{3}$$

The matrix elements are given by

$$M_{\beta\alpha}=\int\int\int_V dV[\mathbf{v}_\beta^*(\mathbf{r})\cdot\mathbf{v}_\alpha(\mathbf{r})], \tag{4}$$

$$\rho N_{\beta\alpha}=\int\int\int_V dV[\lambda(\nabla\cdot\mathbf{v}_\beta(\mathbf{r})^*)(\nabla\cdot\mathbf{v}_\alpha(\mathbf{r})) + 2\mu \mathbf{e}_\beta^*(\mathbf{r}):\mathbf{e}_\alpha(\mathbf{r})], \tag{5}$$

and $S_{\beta\alpha}$ is obtained from

$$\rho S=\lambda_o \int\int_{S_{\hat{\mathbf{n}}}} dS\hat{\mathbf{n}}\cdot\{\mathbf{u}^*(\mathbf{r})\nabla\cdot\mathbf{u}^o(\mathbf{r})\}, \tag{6}$$

where the boundary condition, Eq. (2), has been used. The relation between the gradient of the pressure and the acceleration in the fluid, $\nabla p(\mathbf{r})=-\rho_o \ddot{\mathbf{u}}^o(\mathbf{r})$, and the continuity of the normal component of the displacement at the solid–fluid interface are employed in Sec. IV to evaluate these surface matrix elements. A remark that should be inserted is that the boundary conditions at the solid–fluid interface (or at the solid–vacuum interface) are not imposed on the basis vectors but are imposed on the solution by the determination of the coefficients, A_α.

III. BASIS VECTORS

Our basis vectors for the solid are derived from a separable scalar field

$$\psi_{lmn}(\mathbf{r})=\phi_l^\pm(x)\phi_m^\pm(y)\phi_n^\pm(z), \tag{7}$$

where for the side with length $2a$

$$\phi_o^+(x)=\frac{1}{\sqrt{a}} \quad \text{and} \quad \phi_l^\pm(x)=\frac{1}{2\sqrt{a}}[e^{i\eta_l x}\pm e^{-i\eta_l x}],$$

with $\eta_l=\Theta l/a$ and $-a\leqslant x\leqslant a$ [m indicates the coordinate, $-b\leqslant y\leqslant b$, and n indicates the coordinate, $-c\leqslant z\leqslant c$]. If $\pi\geqslant\Theta$ these functions form a complete set for the volume occupied by the parallelepiped. The $\pm$ sign determines the parity of the scalar function $\phi_l^\pm(-x)=\pm\phi_l^\pm(x)=\varepsilon_{\hat{x}}\phi_l^\pm(x)$, describes its behavior under complex conjugation, $\phi_l^{\pm *}(x)=\pm\phi_l^\pm(x)$, and differentiation, $\partial_x\phi_l^\pm(x)=i\eta_l\phi_l^\mp(x)$. The basis vectors for displacement are formed by applying a gradient operator to the scalar field

$$\mathbf{v}_\alpha=\hat{\mathbf{n}}_\alpha\nu_\alpha=\hat{\mathbf{n}}_\alpha\partial_\alpha\psi_{lmn}(\mathbf{r}), \tag{8}$$

where α denotes the vector direction and the scalar field variables. The coordinates are chosen so that each $\hat{\mathbf{n}}_\alpha$ is normal to a face of the parallelepiped. The choice of basis vectors employed here permits the computation of the pressure on the faces of the parallelepiped using a fast Fourier routine.[8]

Using $\Theta=\pi$ allows one to obtain a convergent solution to the variational procedure for the eigenfrequencies for the solid–vacuum problem; however, the solid–fluid problem also requires the normal displacement at the interfaces. Since the displacements are not periodic, a Fourier series with the edges of the solid as its end points provides a poor representation for them at the interfaces [Gibbs phenomenon]. Therefore, we expand the interval for our representation by taking $\Theta\leqslant\pi$ [we use $\Theta\approx\pi/2$]. This insures that all of the basis vectors do not vanish at an interface and permits an accurate representation of the displacement at the interfaces. It is also the strategy adopted in employing a discrete Fourier transform for a nonperiodic function[8,9] [see Appendix B].

Orthorhombic symmetry permits the displacements to be separated into eight irreducible groups.[13] These groups are characterized by the parity of the components of the displacement vectors. The parity of the scalar field required for each of these groups is displayed in Table I.

TABLE I. Symmetry groupings using the group labelling of Ref. 13. The parity is that of the scalar field.

Group	x parity	y parity	z parity
A_g	+	+	+
B_{1g}	−	−	+
B_{2g}	−	+	−
B_{3g}	+	−	−
A_u	−	−	−
B_{1u}	+	+	−
B_{2u}	+	−	+
B_{3u}	−	+	+

In the nonviscous fluid[12] there is no shear wave and the longitudinal wave in the fluid is described by a scalar field which satisfies the wave equation,

$$(\nabla^2+k^2)\psi_o(\mathbf{r})=0,$$

where $k=\omega/c_o$ is the wave vector and the sound speed in the fluid is c_o. The corresponding Green's function for an outgoing spherical wave in the fluid[6] is

$$g(\mathbf{r}|\mathbf{r}')=e^{ik|\mathbf{r}-\mathbf{r}'|}/4\pi|\mathbf{r}-\mathbf{r}'|, \tag{9}$$

which is employed to evaluate the pressure loading of the parallelepiped due to the fluid.

IV. COMPUTATIONS

The calculation could proceed by solving the matrix equation

$$\Omega M_{\beta\alpha}A_\alpha=(N_{\beta\alpha}-S_{\beta\alpha})A_\alpha, \tag{10}$$

self-consistently for the driving angular frequency, ω_j, such that $\mathrm{Re}\,\Omega_j=\omega_j^2$ [self-consistency is needed since the wave number in the fluid is determined by $k_j=\omega_j/c_o$]. This would also permit the determination of $\tau=\omega/\mathrm{Im}\,\Omega_j$. However, for the ideal gas at the moderate pressures treated here the ratio of the gas density to the solid density, ρ_o/ρ, is much less than 1. This results in only a small change in the frequencies from their vacuum values. For these small changes one can obtain greater accuracy by employing a perturbation approach and solving the free-surface problem with real symmetric matrices

$$\omega_j^2 M_{\beta\alpha} A_\alpha^j = N_{\beta\alpha} A_\alpha^j, \tag{11}$$

for the angular eigenfrequencies, ω_j, of the parallelepiped in a vacuum.

For the driven vibrations at angular frequency, ω_j, Eq. (10) reduces to

$$\omega_j^2 M_{\beta\alpha}\delta A_\alpha^j + \Delta\Omega_j M_{\beta\alpha}(A_\alpha^j+\delta A_\alpha^j) = N_{\beta\alpha}\delta A_\alpha^j - S_{\beta\alpha}(A_\alpha^j+\delta A_\alpha^j).$$

Acting on this with the eigenvector, A_β^{j*}, and, keeping only terms of first order in ρ_o/ρ gives the perturbation of the eigenvalue

$$\Delta\Omega_j = -A_\beta^j{}^* S_{\beta\alpha} A_\alpha^j / A_\beta^j{}^* M_{\beta\alpha} A_\alpha^j. \tag{12}$$

This perturbation can be expressed as

$$\Delta\Omega_j \approx \omega_j(2\Delta\omega_j - i/\tau_j) \propto \rho_o\omega_j^2/\rho, \tag{13}$$

where $\Delta\omega_j/2\pi$ is the shift in the frequency due to gas loading and $1/\omega_j\tau_j$ characterizes the damping due to the radiated energy.

Using $-k^2\rho(\mathbf{r})=\rho_o\omega^2\nabla\cdot\mathbf{u}^o(\mathbf{r})$, which follows from the assumed time dependence, Eq. (6) becomes

$$\rho S = -\frac{\lambda_o}{c_o^2\rho_o}\int\int_{S_{\hat{\mathbf{n}}}} ds\,\hat{\mathbf{n}}\cdot\{\mathbf{u}^*(\mathbf{r})p(\mathbf{r})\}. \tag{14}$$

Hence, in order to evaluate the surface integral matrix elements, $S_{\beta\alpha}$, one needs to determine the pressure on the surface of the parallelepiped due to the outgoing, spherical pressure wave generated by the motion of the surface of the solid. Employing the Green's function for an outgoing spherical wave, Eq. (9), we can express the vibrating pressure in the fluid at $\mathbf{r}$ by[6]

$$p(\mathbf{r}) = \int\int_{S_{\hat{\mathbf{n}}'}} dS'\,\hat{\mathbf{n}}'\cdot\{g(\mathbf{r}|\mathbf{r}')\nabla' p(\mathbf{r}') - p(\mathbf{r}')\nabla' g(\mathbf{r}|\mathbf{r}')\},$$

where the integration is over the surface, $S_{\hat{\mathbf{n}}'}$, of the volume enclosing the fluid and $\hat{\mathbf{n}}'$ is the outward-pointing normal to this surface. Imposing continuity of the normal component of the displacement at the solid–fluid interface, we have for the pressure in the fluid

$$p(\mathbf{r}) = \int\int_{S_{\hat{\mathbf{n}}}} dS'\,\hat{\mathbf{n}}\cdot\{p(\mathbf{r}')\nabla' g(\mathbf{r}|\mathbf{r}')\} = -\rho_o\omega^2\int\int_{S_{\hat{\mathbf{n}}}} dS'\,\hat{\mathbf{n}}\cdot\{g(\mathbf{r}|\mathbf{r}')\mathbf{u}(\mathbf{r}')\}, \tag{15}$$

where the integration is now over the surface of the solid and $\hat{\mathbf{n}}$ is the outward normal to the solid surface. Confining $\mathbf{r}$ to the solid surface with $\hat{\mathbf{n}}\cdot(\mathbf{r}-\mathbf{r}')\geqslant 0^+$ results in a set of inhomogeneous, coupled integral equations which one needs to solve to determine the pressure on the faces of the vibrating solid.

In order to solve the coupled set of equations, the vibrating pressure on a face of the solid is expressed in terms of the divergence of the basis vectors, yielding

$$p(\mathbf{r})|_{S_{\hat{\mathbf{n}}}} = -\sum_\alpha \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}_\alpha P_\alpha \psi_\alpha(\mathbf{r})|_{S_{\hat{\mathbf{n}}}}, \tag{16}$$

on the faces of the solid where α denotes $[l,m,n]$ and the direction of the normal displacement at the face. Using this representation, the surface contribution, Eq. (14), is given by

$$\rho S = \sum_{\beta\alpha} A_\beta^* \int\int_{S_{\hat{\mathbf{n}}}} dS\,\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}_\beta\{\nu_\beta^*(\mathbf{r})\psi_\alpha(\mathbf{r})\}\hat{\mathbf{n}}\cdot\hat{\mathbf{n}}_\alpha P_\alpha, \tag{17}$$

and the integral equations for $p(\mathbf{r})$ are converted to matrix equations. Introducing

$$s_{\beta\alpha} = \int\int_{S_{\hat{\mathbf{n}}_\beta}} dS\,\hat{\mathbf{n}}_\beta\cdot\hat{\mathbf{n}}_\alpha\nu_\beta^*(\mathbf{r})\psi_\alpha(\mathbf{r}), \tag{18}$$

Eq. (17) becomes

$$\rho S = 2\sum_{\beta\alpha} A_\beta^* s_{\beta\alpha} P_\alpha. \tag{19}$$

The factor of 2 accounts for the equality of the contributions from the surfaces parallel and antiparallel to $\hat{\mathbf{n}}_\beta$.

For the surface whose normal is parallel to $\hat{\mathbf{n}}_\beta$, we obtain from Eq. (15)

$$\sum_\alpha (s_{\beta\alpha}-f_{\beta\alpha})P_\alpha = \rho_o\omega_j^2\sum_\alpha g_{\beta\alpha}A_\alpha, \tag{20}$$

where

$$f_{\beta\alpha} = \int\int_{S_{\hat{\mathbf{n}}_\beta}} dS\,\nu_\beta^*(\mathbf{r})\int\int_{S_{\hat{\mathbf{n}}_\alpha}} dS'\,\hat{\mathbf{n}}_\alpha\cdot[\nabla' g(\mathbf{r}|\mathbf{r}') - \varepsilon_{\hat{\mathbf{n}}_\alpha}\nabla' g(\mathbf{r}|\mathbf{r}'_-)]\psi_\alpha(\mathbf{r}'), \tag{21}$$

and

$$g_{\beta\alpha} = \int\int_{S_{\hat{\mathbf{n}}_\beta}} dS\,\nu_\beta^*(\mathbf{r})\int\int_{S_{\hat{\mathbf{n}}_\alpha}} dS'[g(\mathbf{r}|\mathbf{r}') + \varepsilon_{\hat{\mathbf{n}}_\alpha} g(\mathbf{r}|\mathbf{r}'_-)]\nu_\alpha(\mathbf{r}'). \tag{22}$$

Here, the symmetry of the scalar field, $\psi_{lmn}(x,y,-z) = \varepsilon_{\hat{z}}\psi_{lmn}(x,y,z)$, has been used and $\mathbf{r}'_-$ indicates that $\mathbf{r}'$ is to be evaluated on the face antiparallel to $\hat{\mathbf{n}}_\alpha$. Considering the surface whose normal is antiparallel to $\hat{\mathbf{n}}_\beta$ yields an identical matrix equation. The details of the evaluation of these matrix elements are presented in Appendixes A and B. Inverting the matrix multiplying P_α in Eq. (20) and substituting into Eq. (19) results in an expression for the surface matrix elements

$$\rho S_{\beta\alpha} = 2\rho_o\omega_j^2 \sum_{\gamma\alpha} s_{\beta\delta}(s_{\delta\gamma} - f_{\delta\gamma})^{-1} g_{\gamma\alpha}. \tag{23}$$

Considering only the surface at $z=c$, the contribution to the pressure from this face is

$$p(\mathbf{r}) \approx \int\int_{S_{\hat{\mathbf{z}}}} dS' \{ p(\mathbf{r}')\partial_{z'} g(\mathbf{r}|\mathbf{r}') - \rho_o\omega^2 g(\mathbf{r}|\mathbf{r}')\hat{\mathbf{z}}\cdot\mathbf{u}(\mathbf{r}')|_{z'=c}. \} \tag{24}$$

Making use of the Fourier integral transform of the Green's function, one can easily show that

$$\lim_{z\to z'} \partial_{z'} g(\mathbf{r}|\mathbf{r}') = -\frac{1}{2}\mathrm{sgn}(z'-z)\,\delta(x-x')\,\delta(y-y'), \tag{25}$$

and

$$p(\mathbf{r})|_{z=c^+} \approx -2\rho_o\omega^2 \int\int_{S_{\hat{\mathbf{z}}}} dS' \hat{\mathbf{z}}\cdot\{ g(\mathbf{r}|\mathbf{r}')\mathbf{u}(\mathbf{r}')|_{z'=c}\}. \tag{26}$$

This is the zeroth-order approximation (ZO) in an iterative solution of the integral equations for the pressure on the surface of the solid. It is also the result one obtains for the pressure on the surface of a vibrating flat plate set in an infinite baffle [introduction of the baffle permits the construction of a Green's function whose normal derivative vanishes at the surface[14]]. Using this result to approximate the surface matrix elements gives

$$\rho S^o_{\beta\alpha} = 2\rho_o\omega_j^2 \int\int_{S_{\hat{\mathbf{n}}}} dS \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}_\beta \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}_\alpha \times \left\{ \nu^*_\beta(\mathbf{r}) \int\int_{S_{\hat{\mathbf{n}}_\alpha}} dS' [g(\mathbf{r}|\mathbf{r}')\nu_\alpha(\mathbf{r}')] \right\},$$

where $S_{\hat{\mathbf{n}}}$ indicates an integral over the faces of the solid. However, the basis vectors are perpendicular to the faces of the parallelepiped and for each matrix element there are only two contributions from a pair of parallel faces. These contributions are equal; hence

$$\rho S^o_{\beta\alpha} = 4\rho_o\omega_j^2 \hat{\mathbf{n}}_\beta\cdot\hat{\mathbf{n}}_\alpha \int\int_{S_{\hat{\mathbf{n}}_\beta}} dS \nu^*_\beta(\mathbf{r}) \times \int\int_{S_{\hat{\mathbf{n}}_\alpha}} dS' g(\mathbf{r}|\mathbf{r}')\nu_\alpha(\mathbf{r}'). \tag{27}$$

TABLE II. Inverse of the quality factor, Eq. (28), for gas-loaded, fused silica parallelepipeds with dimensions given in Ref. 17. The experimental data and TDM theoretical values are quoted from Ref. 5. The columns headed by theory and ZO (zero order) contain the results of the current calculations. A pressure-independent part has been subtracted from the experimental data, and the three sets of theoretical results only account for the radiation damping due to an ideal gas. The free-surface frequencies are 501.45 (KHz) for the torsional mode and 494.52 (KHz) for the compressional mode.

1×10^4/Qr	Experiment	Theory	ZO	TDM[a]
		Torsional[b]		
Ar	0.599	0.644	0.652	0.962
Air	0.465	0.581	0.594	0.943
He	0.071	0.208	0.209	0.326
		Compressional[c]		
Ar	1.053	0.607	0.600	0.962[d]
He	0.318	0.199	0.197	0.336

[a]Zhang *et al.*, Ref. 5.
[b]Sample 1.
[c]Sample 2.
[d]TDM calculation using dimensions of sample 1.

This approximation is similar to the TDM; however, it incorporates an accurate description of the normal displacements at the surface of the solid.

V. RESULTS

Our calculations are for a solid of fused silica with material parameters[15]

$$\rho = 2.2185\times10^3 \ \ (\mathrm{kg/m^3}), \quad \mu = 0.3117\times10^6 \ \ (\mathrm{bar}),$$

and

$$\lambda = 0.1597\times10^6 (\mathrm{bar}),$$

for which there are experimental RUS results for gas-loaded samples [parallelepipeds[5] and sphere[7,16]]. The lengths of the sides of the parallelepiped are $2a$, $2b$, and $2c$, and, introducing the aspect ratios, $r_y = b/a$ and $r_z = c/a$, so that

$$r_o = \sqrt{a^2+b^2+c^2} = a\sqrt{1+r_y^2+r_z^2},$$

the eigenfrequencies scale with a. The scaled angular eigenfrequencies, $a\omega_j$, for the free-surface vibrations are functions of r_y and r_z, and the material parameters of the solid. A nonviscous ideal gas at a fixed temperature can be characterized by two independent quantities: the density, ρ_o, and the sound speed, c_o; the perturbation of the eigenvalue, Eq. (13), is proportional to the density ratio ρ_o/ρ. The experimental data can be characterized by the quality factor for the radiation damping of the mode

$$Q_r = \omega_j^2/\mathrm{Im}\,\Delta\Omega_j = \omega_j\tau_j, \tag{28}$$

and the ratio of the frequency shift to the eigenfrequency, and these quantities depend on the static pressure through the density ratio.

The experimental data[5] are for the lowest torsional mode, A_{u1}, in sample 1 (Ref. 17) and the lowest compressional mode, A_{g1}, in sample 2.[17] In Tables II and III the experimental results[5] for these modes are compared to our theoretical computations and the results from the TDM

TABLE III. Frequency shift with gas loading of fused silica parallelepipeds with dimensions given in Ref. 17. The experimental data and TDM theoretical values are quoted from Ref. 5. The columns headed by Theory and ZO (zero order) contain the results of the current calculations. The three sets of theoretical results only account for the shift due to the radiation damping of an ideal gas, while the experimental data are also dependent on the change in moduli of the samples. The free-surface frequencies are 501.45 (KHz) for the torsional mode and 494.52 (KHz) for the compressional mode.

$\Delta f \times 10^6/f$	Experiment	Theory	ZO	TDM[a]
		Torsional[b]		
Ar	−13.4	−7.25	−4.32	−3.31
Air	−46.2	−5.08	−3.40	−3.74
He	−10.5	−4.64	−3.11	−4.14
		Compressional[b]		
Ar		−0.02	−0.50	−0.02
He		−0.21	−0.44	−0.03
		Compressional[c]		
Ar	9.37	−0.90	−0.76	
He	−6.19	−1.07	−0.92	

[a]Zhang *et al.*, Ref. 5.
[b]Sample 1.
[c]Sample 2.

calculation[5] [Note: the compressional calculation for the TDM was performed using the dimensions of sample 1.] The three theoretical calculations only account for the radiation damping by an ideal gas [no viscosity and no allowance for the change in the elastic moduli of the solid with pressure].

The values of $1/Q_r$ are reported in Table II. The experimental values are from a fit of the resonant peaks and they have a pressure independent factor $1/Q_o$ subtracted from the total.[5] Overall, there is good agreement between the theoretical and experiment results for this factor. The experimental frequency shifts reported in Table III are much larger than the theoretical shifts. Since the experimental results also include the shifts due to the change in the elastic moduli with pressure, differences are expected. However, there is no pattern to the differences which would permit the determination of the intrinsic pressure dependence of the elastic moduli from these data.

The comparison of the TDM with the current calculations indicates the good predictive strength of this simple mechanical model for the gas loading. It gives the correct magnitudes for the quality factors and the frequency shifts, and a comparison of Figs. 14 and 15 of Ref. 5 with our Figs. 1 and 2 shows that it provides the correct overall wave number dependence of the acoustical radiation impedance. Any detailed comparison should be between the TDM and the ZO calculations, since they each employ a vibrating surface in an infinite baffle to describe the surface pressure. However, the TDM introduces a surface velocity distribution and employs a mechanical model to combine the contributions to the impedance from the faces of the parallelepiped, while in the ZO the actual surface velocity of the vibrating solid is used and the contributions are weighted by an integral over the volume.

In Figs. 1 and 2, the acoustic radiation resistance, $R(k)$, and radiation reactance, $X(k)$, are plotted for the current computations. These are functions of the geometry of the solid and wave number, and they are defined by

$$R(k) - iX(k) = \frac{ik\rho}{4\rho_o} \frac{\Delta\Omega_j}{\omega_j^2}. \tag{29}$$

In all cases the results are displayed for the lowest frequency mode of each symmetry, and the aspect ratios are chosen so that $r_y = r_z$. Further, the sizes of the parallelepipeds are chosen so that the modes have a frequency of 500 (KHz). The wave number is proportional to the reciprocal of the sound speed in the gas, and the wave numbers for helium and argon at 17 °C are indicated on the plots.

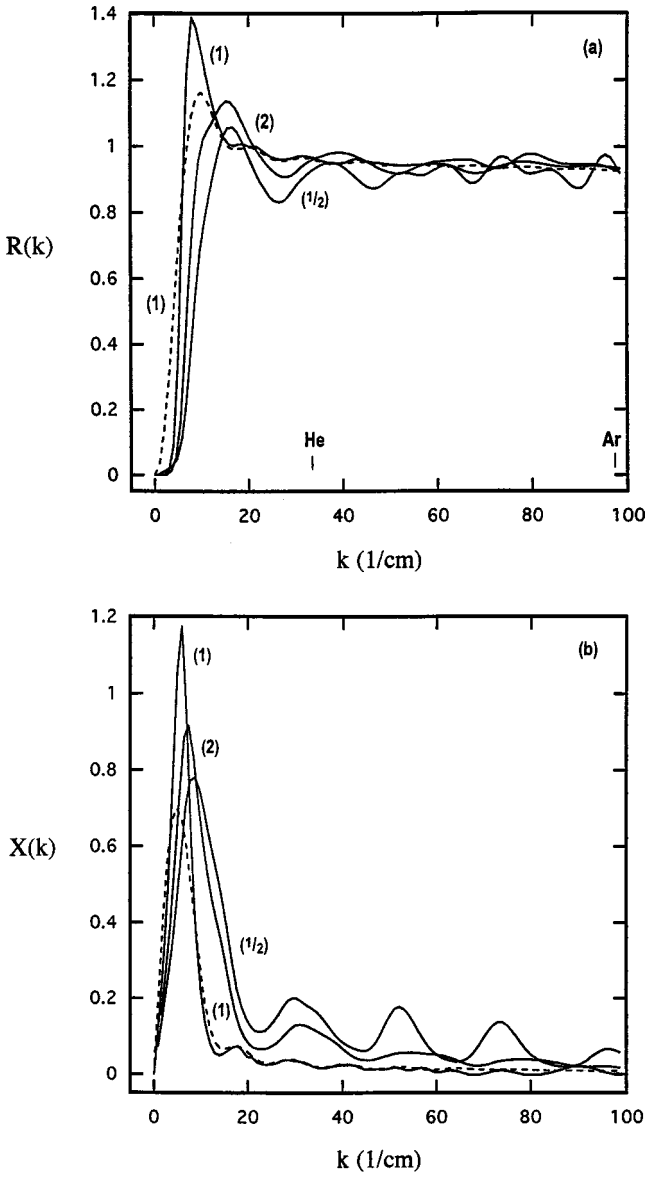


FIG. 1. Acoustic (a) radiation resistance, $R(k)$, and (b) radiation reactance, $X(k)$, versus wave number for the compressional mode, A_{g1}, of rectangular parallelepipeds in a pressurizing ideal fluid. The curves are labeled by the aspect ratios for the parallelepiped [$r_y = r_z = 1/2, a = 0.2842$ (cm); $r_y = r_z = 1, a = 0.2650$ (cm); $r_y = r_z = 2, a = 0.1326$ (cm)]. The free-surface vibrational frequency is 500 (KHz) which is the lowest frequency mode for all the parallelepipeds. The dashed lines are the ZO (zero order) results and the solid lines are the full perturbation results. The wave number is inversely proportional to the sound speed in the fluid and the values for helium and argon are indicated on the axis.

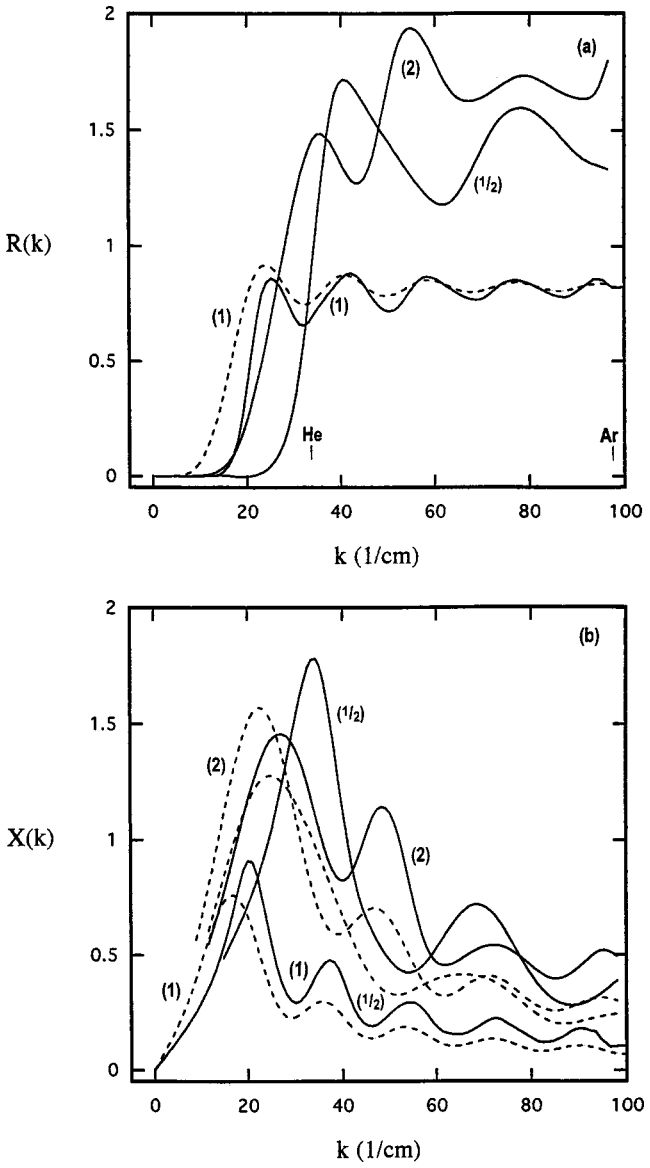


FIG. 2. Acoustic (a) radiation resistance, $R(k)$, and (b) radiation reactance, $X(k)$, versus wave number for the torsional mode, A_{u1}, of rectangular parallelepipeds in a pressurizing ideal fluid. The curves are labeled by the aspect ratios for the parallelepiped $[r_y=r_z=1/2, a=0.1724\ (\mathrm{cm}); r_y=r_z=1, a=0.1705\ (\mathrm{cm}); r_y=r_z=2, a=0.0640\ (\mathrm{cm})]$. The free-surface vibrational frequency is 500 (KHz), which is the lowest frequency mode for all the parallelepipeds. The dashed lines are the ZO (zero order) results and the solid lines are the full perturbation results. The wave number is inversely proportional to the sound speed in the fluid, and the values for helium and argon are indicated on the axis.

The correspondence in Fig. 1 between the ZO and the full calculations of the acoustical impedance for the compressional mode reflects the small displacement of the parallelepiped edges relative to the displacement of the center of the faces for this mode. These small displacements permit an accurate description of the vibrations by employing baffled surfaces [see Fig. 13 and arguments in Ref. 5]. These small displacements can also account for the very small dependence on shape of the short wavelength impedance for this mode [note that the largest dimension of all three parallelepipeds is ~0.55 (cm) and the wavelength at this frequency in argon is ~0.065 (cm)]. The small reactance at large wave number and small shape dependence of this mode argue for its usefulness in determining the effect of pressure on the elastic moduli.

The calculated results for the torsional mode are presented in Fig. 2, and they reveal the strong shape dependence of the acoustical impedance for this mode where the large displacements are at the edges of the parallelepiped. The suppression of the resistance for the long wavelength vibrations due to the quadrupole nature of the vibrating surface is clearly seen in Fig. 2(a) as it was in the TDM calculation [see Fig. 13 and arguments in Ref. 5]. The average magnitude of the short wavelength resistance of the full perturbation calculation for all three shapes closely matches that found for the ZO calculation. However, the amplitude of the oscillatory portion of the resistance and the reactance of this mode is much larger for the full calculation, which reflects the coupling of the neighboring faces due to the relatively large displacement of the edges of the faces.

VI. CONCLUDING REMARKS

Since the ZO calculation uses only the pressure on oscillating, baffled surfaces, it can easily be applied to the computation of the effect of gas loading on the vibrational spectrum of single-crystal samples. Utilizing Eqs. (12) and (27), we can write the perturbation of the eigenvalue for this approximation as

$$\Delta\Omega_j^o = 2\omega_j^2\rho_o \int\int_{S_{\hat{\mathbf{n}}}} dS\,\hat{\mathbf{n}}\cdot\mathbf{u}^*(\mathbf{r}) \int\int_{S_{\hat{\mathbf{n}}}} dS'\,g(\mathbf{r}|\mathbf{r}') \times \hat{\mathbf{n}}\cdot\mathbf{u}(\mathbf{r}') \Big/ \rho \int\int\int_V dV\,\mathbf{u}^*(\mathbf{r})\cdot\mathbf{u}(\mathbf{r}), \qquad (30)$$

where the displacements are for the mode with angular eigenfrequency ω_j. These displacements can easily be assembled using the eigenvectors, A_α^j, which satisfy Eq. (11) and the three, two-dimensional integrals with Green's function kernels can be evaluated numerically. A less sophisticated estimate can be obtained by assuming a description of the displacements for a particular mode as was done for the TDM. The integral in the numerator of Eq. (30) is the same integral needed for the TDM estimate.[5] However, for the TDM estimate this integral over each face was divided by an integral over the displacement on that face, and these contributions were averaged to give an estimate for the solid. Hence, one only needed to assume a form for the surface displacements—no relative amplitudes were needed. In the ZO estimate the surface terms are combined and then divided by an average over the solid; hence, an assumption of the displacements throughout the solid is required.[18]

For the compressional mode, A_{g1}, the frequency shift due to gas loading of the solid is small, particularly for the large wave number response of the gas. Thus, the analysis of experimental data for this mode permits a direct estimate of the magnitude of the pressure dependence of the elastic moduli, and the ZO approximation provides an accurate estimate of the contribution from the gas loading.

The formalism clearly does not depend on the set of basis functions and is not limited to parallelepipeds. The sur-

face displacement (or velocity) can be constructed from the Ritz–Rayleigh solution of the free-surface problem of the vibrating solid, Eq. (11), and the evaluation of the pressure using this displacement can then be treated as an independent problem. The ZO approximation for the pressure may suffice for some applications. For applications where greater accuracy is required, the coupled, inhomogeneous integral equations, Eq. (15), for the pressure on the faces of the solid need to be solved, and the full calculation reported here provides a solution. The pressure could also be numerically evaluated by employing the well-studied boundary element method[19] or seeking an iterative solution of Eq. (15).

ACKNOWLEDGMENTS

The author wishes to acknowledge discussions with M. Levy and R.S. Sorbello. Further, the correspondence and critical comments of D.G. Isaak have been very beneficial to the author in the preparation of this manuscript.

APPENDIX A: INTEGRALS FOR THE SURFACE MATRIX ELEMENTS

In this Appendix the integrals required for the surface matrix elements in our Rayleigh–Ritz calculation for the parallelepiped are summarized and reduced.

The matrix elements $s_{\beta\alpha}$, Eq. (18), are easily assembled using the overlap integrals for the basis functions

$$O_{l'l}=\int_{-a}^{a} dx\,\phi_{l'}^{*}(x)\,\phi_{l}(x) =S((\eta_l-\eta_{l'})a)+\varepsilon_x S((\eta_l+\eta_{l'})a), \tag{A1}$$

where $S(q)=\sin(q)/q$.

The difficult integrals to evaluate are $f_{\beta\alpha}$ and $g_{\beta\alpha}$, Eqs. (21) and (22). For $f_{\beta\alpha}$ on a particular face there are three contributions to consider. Setting $\hat{\mathbf{n}}_\beta=\hat{\mathbf{z}}$ and considering the pressure on the face at $z=c$. From the faces at $z=c$ and $z=-c$

$$f_{\beta\alpha}^{c}=\int\int dxdy\,\nu_{\beta}^{*}(x,y,c)\int\int dx'dy'\{\psi_\alpha(x',y',z') \times\partial_{z'}g(x,y,c^{+}|x',y'z')|_{z'=c} -\psi_\alpha(x',y',z')\partial_{z'}g(x,y,c^{+}|x',y',z')|_{z'=-c}\},$$

and employing Eq. (25)

$$f_{\beta\alpha}^{c}=\int\int dxdy\,\nu_{\beta}^{*}(x,y,c)\left\{\frac{1}{2}\psi_\alpha(x,y,c) -\int\int dx'dy'\,\psi_\alpha(x',y',z') \times\partial_{z'}g(x,y,c^{+}|x',y',z')|_{z'=-c}\right\},$$

or

$$f_{\beta\alpha}^{c}=\frac{1}{2}s_{\beta\alpha}+\varepsilon_{\hat{z}}\int\int dxdy\,\nu_{\beta}^{*}(x,y,c) \times\int\int dx'dy'\,\psi_\alpha(x',y',z') \times\partial_{z'}g(x,y,-c^{+}|x',y',z')|_{z'=c}. \tag{A2}$$

From the faces at $x=a$ and $x=-a$, the contribution is

$$f_{\beta\alpha}^{a}=\int\int dxdy\,\nu_{\beta}^{*}(x,y,c)\int\int dy'dz' \times\partial_{x'}g(x,y,c^{+}|x',y',z')|_{x'=a^{+}}\{\psi_\alpha(x',y',z')|_{x'=a^{+}} +\varepsilon_{\hat{x}}\psi_\alpha(x',y',z')|_{x'=-a^{+}}\},$$

since $|x|\leqslant a$ and $\partial_{x'}g(x,y,c^{+}|x',y',z')|_{x'=-a^{+}} =-\partial_{x'}g(-x,y,c^{+}|x',y',z')|_{x'=a^{+}}$. This reduces to

$$f_{\beta\alpha}^{a}=2\int\int dxdy\,\nu_{\beta}^{*}(x,y,c)\int\int dy'dz' \times\partial_{x'}g(x,y,c^{+}|x',y',z')\psi_\alpha(x',y',z')|_{x'=a^{+}}.$$

Switching the derivative of the Green's function from x' to x and integrating by parts yields

$$f_{\beta\alpha}^{a}=2\int\int dxdy\,\partial_x\nu_{\beta}^{*}(x,y,c)\int\int dy'dz' \times g(x,y,c^{+}|a^{+},y',z')\psi_\alpha(a^{+},y',z') -2\int dy\,\nu_{\beta}^{*}(x,y,c)\int\int dy'dz' \times g(x,y,c^{+}|a^{+},y',z')\psi_\alpha(a^{+},y',z')|_{x=-a}^{x=a}. \tag{A3}$$

The contribution from the faces at $y=b$ and $y=-b$, and the matrix elements for the other directions of $\hat{\mathbf{n}}_\beta$ are easily obtained from the symmetry by permutation of the coordinates.

Considering $g_{\beta\alpha}$ for $\hat{\mathbf{n}}_\beta=\hat{\mathbf{z}}$; for $\hat{\mathbf{n}}_\alpha=\hat{\mathbf{z}}$

$$g_{\beta\alpha}^{c}=\int\int dxdy\,\nu_{\beta}^{*}(x,y,c)\int\int dx'dy'\,\nu_\alpha(x',y',c) \times\{g(x,y,c^{+}|x',y',c)+\varepsilon_{\hat{z}}g(x,y,c^{+}|x',y',-c)\}, \tag{A4}$$

and for $\hat{\mathbf{n}}_\alpha=\hat{\mathbf{x}}$

$$g_{\beta\alpha}^{a}=\int\int dxdy\,\nu_{\beta}^{*}(x,y,c)\int\int dy'dz'\,\nu_\alpha(a,y',z') \times\{g(x,y,c^{+}|a^{+},y',z') +\varepsilon_{\hat{x}}g(x,y,c^{+}|-a^{+},y',z')\} =2\int\int dxdy\,\nu_{\beta}^{*}(x,y,c)\int\int dy'dz'\,\nu_\alpha(a,y',z') \times g(x,y,c^{+}|a^{+},y',z'), \tag{A5}$$

with a corresponding expression for $\hat{\mathbf{n}}_\alpha=\hat{\mathbf{y}}$. Hence, the evaluation of these matrix elements requires the integrals having the forms

$$D^{c}_{l'm':lm}=\int\int_{S_{\hat{\mathbf{z}}}} dS'\,\phi^{*}_{m'}(y')\phi^{*}_{l'}(x') \times\int\int_{S_{\hat{\mathbf{z}}}} dSg(x',y',c^{+}|x,y,z)\phi_{l}(x)\phi_{m}(y), \tag{A6}$$

and

$$C^{c,a}_{l'm':mn}=\int\int_{S_{\hat{\mathbf{z}}}} dS'\,\phi^{*}_{m'}(y')\phi^{*}_{l'}(x') \times\int\int_{S_{\hat{\mathbf{x}}}} dSg(x',y',c^{+}|a^{+},y,z)\phi_{m}(y)\phi_{n}(z). \tag{A7}$$

Noting the dependence of the Green's function, Eq. (9), on the coordinates

$$|\mathbf{r}'-\mathbf{r}|=\sqrt{(x'-x)^2+(y'-y)^2+(z'-z)^2},$$

consider

$$I_l=\int_{-a}^{a} dx\,\phi_l(x)f(|a^{+}-x|) =-\frac{i\eta_l}{2}\int_0^{2a} dsf(|s|)[S(\eta_l a)(e^{-i\eta_l s}-\varepsilon_{\hat{\mathbf{x}}}e^{i\eta_l s}) +iC(\eta_l a)(e^{-i\eta_l s}+\varepsilon_{\hat{\mathbf{x}}}e^{i\eta_l s})], \tag{A8}$$

where $C(q)=\cos(q)/q$. The other integral to be considered is

$$J_{l'l}=\int_{-a}^{a} dx\int_{-a}^{a} dx'\,\phi^{*}_{l'}(x')\phi_l(x)f(|x'-x|), \tag{A9}$$

where this integral vanishes if the basis functions do not have the same parity. Employing the parity, rearranging and changing variables, gives

$$J_{l'l}=2\int_0^{2a} dsf(|s|)\int_{s-a}^{a} dx\,\phi^{*}_{l'}(x-s)\phi_l(x),$$

which for $l\neq l'$ integrates to

$$J_{l'l}=\frac{1}{2}\int_0^{2a} dsf(|s|)\{O_{l'l}(e^{i\eta_l s}+e^{-i\eta_l s}+e^{i\eta_{l'} s} +e^{-i\eta_{l'} s})+iQ_{l'l}(e^{i\eta_l s}-e^{i\eta_l s}) +iQ_{ll'}(e^{i\eta_{l'} s}-e^{-i\eta_{l'} s})\}, \tag{A10a}$$

where

$$Q_{l'l}=C((\eta_l-\eta_{l'})a)+\varepsilon_{\hat{\mathbf{x}}}C((\eta_l+\eta_{l'})a).$$

For $l=l'$ and nonzero

$$J_{ll}=\frac{1}{2}\int_0^{2a} dsf(|s|)\left\{\left(2O_{ll}-\frac{s}{a}\right)(e^{i\eta_l s}+e^{-i\eta_l s}) +2i\varepsilon_{\hat{\mathbf{x}}}C(2\eta_l a)(e^{i\eta_l s}-e^{-i\eta_l s})\right\}, \tag{A10b}$$

and for $l=l'=0$

$$J_{00}=2\int_0^{2a} dsf(|s|)\left(2-\frac{s}{a}\right). \tag{A10c}$$

These expressions, Eqs. (A8) and (A10), permit the integrals, Eqs. (A5) and (A6), to be assembled using the two-dimensional integrals

$$Z_{lm}=Z^{a}_{lm}+Z^{b}_{lm}=\frac{1}{4\pi}\int_0^{2a} dx\int_0^{2b} dye^{i(\eta_l x+\eta_m y)}\{e^{ikr}/r\}, \tag{A11}$$

and the three-dimensional integrals

$$Z_{lmn}=Z^{a}_{lmn}+Z^{b}_{lmn}+Z^{c}_{lmn} =\frac{1}{4\pi}\int_0^{2a} dx\int_0^{2b} dy\int_0^{2c} dze^{i(\eta_l x+\eta_m y+\eta_n z)}\{e^{ikr}/r\}, \tag{A12}$$

where $r=\sqrt{x^2+y^2+z^2}$. The three-dimensional integral has a singularity at the origin. However, by partitioning this volume integral into integrals over three wedge-shaped pieces, one obtains integrals with nonsingular integrands. The piece terminated by the face at $x=2a$ and delineated by the lines joining the origin to the upper corners of this face is

$$Z^{a}_{lmn}=\frac{1}{4\pi}\int_0^{2a} dx\int_0^{bx/a} dy\int_0^{cx/a} dze^{i(\eta_l x+\eta_m y+\eta_n z)} \times\{e^{ikr}/r\}. \tag{A13}$$

Introducing the new variables, $x'=x/a$, $s=ay/bx$, and $t=az/cx$, gives

$$Z^{a}_{lmn}=\frac{abc}{4\pi}\int_0^1 ds\int_0^1 dt\int_0^2 x'dx' \times\{\exp[ix'(k\sqrt{a^2+b^2s^2+c^2t^2}+\Theta l+\Theta ms +\Theta nt)]/\sqrt{a^2+b^2s^2+c^2t^2}\}.$$

The integrand has no singularity and the integration over x' can be performed analytically, leaving a two-dimensional integral to be evaluated numerically. The other two pieces of the integral are obtained by the interchange of the paired variables (a,l) and (b,m), or (a,l) and (c,n). For the two-dimensional surface integrals with a singular integrand [Eq. (A11) with $z=0$] the integral is partitioned into two integrals over triangular-shaped pieces. Introducing new variables again allows one to eliminate the singularity and perform one integration analytically.

Noting that $\partial_x e^{ikr}=ikxe^{ikr}/r$, the integral

$$X_{lmn}=\frac{1}{4\pi ik}\int_0^{2a} dx\int_0^{2b} dy\int_0^{2c} dze^{i(\eta_l x+\eta_m y+\eta_n z)}e^{ikr}, \tag{A14}$$

and a number of two-dimensional integrals are also used. The numerical cost of the direct evaluation of all the two-dimensional integrals for a set of $[l,m,n]$ values is quite high. However, the numerical algorithms for discrete Fourier transforms (DFT) are very efficient and a fast Fourier routine is employed to numerically evaluate these integrals. The procedure for the setup of the DFT and the treatment of the integrable singularity are set forth in Appendix B.

APPENDIX B: DISCRETE FOURIER TRANSFORM EVALUATION OF INTEGRALS

The numerical algorithms for discrete Fourier transforms (DFT) are very efficient,[20] and a fast Fourier routine[8] is employed to numerically evaluate the integrals needed for the surface matrix elements. The discrete Fourier transform[9] is given by

$$F_k = \sum_{j=0}^{N-1} w^{kj} f_j , \tag{B1}$$

where $w = e^{2\pi i/N}$ and the imposed periodicity is such that $f_0 = f_N$. Making use of a Simpson's rule expression for the integral gives

$$\int_0^{2a} dx f(x) e^{i\eta_l x} = \frac{2a}{M} \sum_{j=0}^{M} (e^{2\Theta i/M})^{lj} f_j , \tag{B2}$$

where $x_j = 2aj/M$, $f_0 = \frac{1}{2} f(0)$, $f_M = \frac{1}{2} f(2a)$, and $f_j = f(x_j)$ for $j = 1,...,M-1$. To match the discrete Fourier expression one sets $\Theta = \pi M/N$ and pads the series with zeros; i.e., $f_j = 0$ for $j = M+1,...,N-1$. Using a multidimensional DFT, the integrals with nonsingular integrands can be directly evaluated.

In order to eliminate the integrable singularity of an integral, one employs the shift theorem for DFT[21] to write

$$G_k = F_k - F_{k-1} = \sum_{j=0}^{N-1} w^{kj} [1 - w^{-j}] f_j . \tag{B3}$$

The DFT routine outputs the positive wave number transforms, G_k, and the negative wave number transforms, $G_{-k} = G_{N-k}$, for $k = 1,...,(N-1)/2$, and the periodicity dictates that $G_{-N/2} = G_{N/2}$. To unfold the shifted transform input, F_0, from a separate numerical calculation and obtain

$$F_k = F_0 + \sum_{j=1}^{k} G_j \quad \text{and} \quad F_{-k} = F_0 - \sum_{j=N-k+1}^{N} G_j , \tag{B4}$$

since $G_N = G_0$ and $F_N = F_0$. These are the transforms for $k = 1,...,N/2$, and $F_{-N/2} \neq F_{N/2}$ if $g_0 \neq 0$. In general, for an integrand with an integrable singularity, one expects F_0 to be large compared to the other transforms, since the integral contains no oscillatory exponential factor. This procedure has the advantage of eliminating this term from the DFT calculation and reintroducing it as an addition to all the other transforms.

Employing this theorem, Eq. (B3), for a three dimensional integral, Eq. (A12), start with the one-dimensional DFT

$$H_{00n} = Z_{00n} - Z_{00,n-1} = \sum_{ijk} w_z^{nk} [1 - w_z^{-k}] z_{ijk} ,$$

which is nonsingular, since

$$\lim_{n \to 0} [1 - w_z^{-k}] z_{00n} = abi/4 M_x M_y N_z ,$$

for the Green's function kernel. Then, evaluating Z_{000} by numerical integration of the two-dimensional integrals obtained from Eq. (A13), Eq. (B4) can be used to generate all of the three-dimensional integrals with $l = m = 0$. Employing the theorem for the two-dimensional DFT

$$H_{0mn} = Z_{0mn} - Z_{0,m-1,n} = \sum_{ijk} w_y^{mj} w_z^{nk} [1 - w_y^{-j}] z_{ijk} ,$$

gives all the integrals with $l = 0$. Finally, the three-dimensional DFT

$$H_{lmn} = Z_{lmn} - Z_{l-1,mn} = \sum_{ijk} w_x^{li} w_y^{mj} w_z^{nk} [1 - w_x^{-i}] z_{ijk} ,$$

permits the generation of all the three-dimensional integrals. The first two steps of this procedure suffice for the evaluation of the two-dimensional integrals, Eq. (A11), with an integrable singularity.

[1] H. H. Demarest, Jr., "Cube-resonance method to determine the elastic constants of solids," J. Acoust. Soc. Am. **49**, 768–775 (1971); I. Ohno, "Free vibration of a rectangular parallelepiped crystal and its application to determination of elastic constants of orthorhombic crystals," J. Phys. Earth **24**, 355–379 (1976); A. Migliori and T. W. Darling, "Resonant ultrasound spectroscopy for materials studies and nondestructive testing," Ultrasonics **34**, 473–476 (1996); Albert Migliori and John L. Sarrao, *Resonant Ultrasound Spectroscopy* (Wiley, New York, 1997).

[2] 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 elastic moduli of solids," Physica B **183**, 1–24 (1993).

[3] D. G. Isaak, O. L. Anderson, and T. Coto, "Measured elastic moduli of single-crystal MgO up to 1800 K," Phys. Chem. Miner. **16**, 704–773 (1989); D. G. Isaak, "High temperature elasticity of iron-bearing olivines," J. Geophys. Res. **97**, 1871–1885 (1992); O. L. Anderson and D. G. Isaak, in *Mineral Physics and Crystallography: A Handbook of Physical Constants*, AGU Reference Shelf 2 (American Geophysical Union, Washington, DC, 1995), pp. 64–97.

[4] J. Maynard, "Resonant ultrasound spectroscopy," Phys. Today **49**, 26–31 (1996).

[5] H. Zhang, R. S. Sorbello, C. Hucho, J. Herro, J. R. Feller, D. E. Beck, M. Levy, D. Isaak, J. D. Carnes, and O. Anderson, "Radiation impedance of resonant ultrasonic spectroscopy modes in fused silica," J. Acoust. Soc. Am. **103**, 2385–2394 (1998).

[6] P. M. Morse and K. Uno Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), pp. 319–321.

[7] R. S. Sorbello, J. Feller, M. Levy, D. G. Isaak, J. D. Carnes, and O. L. Anderson, "The effect of gas loading on the RUS spectra of spheres," J. Acoust. Soc. Am. **107**, 808–818 (2000). The torsional modes in spheres only couple to the pressurizing gas through the viscosity of the gas while the compressional modes couple directly to the vibrating pressure wave. As a result of the weak coupling of the torsional modes, the distortion of the spectra of these modes is at least an order of magnitude smaller than for the compressional modes in fused silica spheres. The torsional mode in a parallelepiped couples directly to the pressure wave.

[8] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge & New York, 1989), pp. 386–395.

[9] V. Cizek, *Discrete Fourier Transforms and Their Applications* (Hilger, Bristol, 1986), Chap. 4.

[10] P. M. Morse and H. Feshbach, *Methods of Mathematical Physics* (McGraw-Hill, New York, 1953), Vol. I, pp. 298–301.

[11] L. Brekhovskikh and V. Goncharov, *Mechanics of Continua and Wave*

Dynamics (Springer, Berlin, Heidelberg, 1985), pp. 48–50.
[12] The Navier–Stokes equation describes the behavior of a viscous fluid and, relating the pressure to the displacement, one obtains an equation of the same form as that for an isotropic solid [See Ref. 6, p. 279].
[13] E. Mochizuki, ''Application of group theory to free oscillations of an anisotropic rectangular parallelepiped,'' J. Phys. Earth **35**, 159–170 (1987).
[14] Reference 6, pp. 366–369.
[15] I. S. Grigoriev and E. Z. Meilikhov, *Handbook of Physical Quantities* (CRC Press, Boca Raton, 1997), p. 178. Fused quartz: longitudinal speed $=5.935\times10^3$ (m/s); shear speed $=3.740\times10^3$ (m/s).
[16] D. G. Isaak, J. D. Carnes, O. L. Anderson, and H. Oda, ''Elasticity of fused silica spheres under pressure using resonant ultrasound spectroscopy,'' J. Acoust. Soc. Am. **104**, 2200–2206 (1998).
[17] Zhang *et al.* [Ref. 5] quote experimental data for two fused silica samples. Sample 1: $0.3194\times0.3083\times0.2491$ (cm^3) and mass 0.0545 (g) was used to obtain the torsional mode data. Sample 2: $0.569\times0.426\times0.318$ (cm^3) (same density as sample 1) was used to obtain the compressional mode data [These are the nominal dimensions of this sample, they were not quoted in Ref. 5].
[18] If one considers a cube, the ZO, Eq. (30), has six identical terms in the numerator and three identical terms in the denominator. Using the cosine surface displacement introduced in Ref. 5, the TDM and the ZO estimates give the same answers for the cube.
[19] One of the reviewers of this manuscript indicated the applicability of the boundary element method to this determination. There is a vast literature on this numerical method including computer programs for its implementation. A short introduction to the method is found in the text by Zhou, *Numerical Analysis of Electromagnetic Fields* (Springer, Berlin, 1993), Chap. 9, and the monogram by Ihlenburg, *Finite Element Analysis of Acoustic Scattering* (Springer, New York, 1998), provides a careful analysis of finite-element computation for a radiating solid.
[20] E. G. Williams and J. D. Maynard, ''Numerical evaluation of the Rayleigh integral for planar radiators using the FFT,'' J. Acoust. Soc. Am. **72**, 2020–2030 (1982). These authors use the fast Fourier transform algorithm for baffled, planar radiators and show that it increases the speed of computation over numerical integration by at least a factor of 400. Although they afford the singularity special treatment, their procedure does not eliminate it from the transform.
[21] Reference 9, pp. 62–63.

A new method to predict the evolution of the power spectral density for a finite-amplitude sound wave

Penelope Menounou
Department of Mechanical Engineering, The University of Texas at Austin, 1 University Station C2200, Austin, Texas 78712-0292

David T. Blackstock
Department of Mechanical Engineering and Applied Research Labortories, The University of Texas at Austin, 1 University Station C2200, Austin, TX 78712-0292

(Received 30 March 1999; revised 17 October 2003; accepted 10 November 2003)

A method to predict the effect of nonlinearity on the power spectral density of a plane wave traveling in a thermoviscous fluid is presented. As opposed to time-domain methods, the method presented here is based directly on the power spectral density of the signal, not the signal itself. The Burgers equation is employed for the mathematical description of the combined effects of nonlinearity and dissipation. The Burgers equation is transformed into an infinite set of linear equations that describe the evolution of the joint moments of the signal. A method for solving this system of equations is presented. Only a finite number of equations is appropriately selected and solved by numerical means. For the method to be applied all appropriate joint moments must be known at the source. If the source condition has Gaussian characteristics (it is a Gaussian noise signal or a Gaussian stationary and ergodic stochastic process), then all the joint moments can be computed from the power spectral density of the signal at the source. Numerical results from the presented method are shown to be in good agreement with known analytical solutions in the preshock region for two benchmark cases: (i) sinusoidal source signal and (ii) a Gaussian stochastic process as the source condition.
[DOI: 10.1121/1.1639902]

PACS numbers: 43.25.Cb, 43.25.Ed [LCS] Pages: 567–580

I. INTRODUCTION

The present paper is a theoretical investigation of the effects of nonlinearity and dissipation on the power spectral density of a progressive sound wave. Traditionally, the signal is known at the source, and time-domain methods can be used to predict the evolution of the signal as it propagates. In many case, however, the time history of the signal is not known at the source. Only its power spectral density (PSD) is given. Consider for example the case of a stochastic process as the source condition. A Gaussian stochastic process with zero mean is characterized by its PSD. The time signal at the source is not given, nor can it be found from the PSD of the process. One can also consider cases of deterministic signals, where the PSD of the signal is known but not the signal itself. In aircraft-noise measurements, for example, the measured signal itself is not reported; the PSD is given instead. The PSD does not contain any phase information, and thus the source signal time history cannot be found from the reported PSD. A method is needed that can be used to predict the evolution of the PSD of a stochastic process or a deterministic signal, when only the PSD of the process or the signal is known at the source.

The Burgers equation is employed for the mathematical description of the combined effects of nonlinearity and dissipation

$$\frac{\partial p}{\partial x}=\frac{\beta}{2\rho_0 c_0^3}\frac{\partial p^2}{\partial t}+\frac{\delta}{2c_0^3}\frac{\partial^2 p}{\partial t^2}, \tag{1}$$

where p is the sound pressure, x is the coordinate along which the plane wave propagates, t is the retarded time $t=t'-x/c_0$ associated with the propagation in the $x+$ direction, c_0 is the small-signal sound speed, ρ_0 is the ambient density of the fluid, β is the dimensionless coefficient of nonlinearity, and δ is the diffusivity for a thermoviscous fluid[1]

$$\delta=\frac{1}{\rho}\left[\frac{4}{3}\mu+\mu_B+\kappa\left(\frac{1}{c_v}-\frac{1}{c_p}\right)\right], \tag{2}$$

expressed as a function of the fluid's shear viscosity μ, bulk viscosity μ_B, thermal conductivity κ, specific heat at constant pressure c_p, and specific heat at constant volume c_v.

For the theoretical analysis as well as for the numerical implementation, it is very helpful to express the Burgers equation (BE) in dimensionless variables. We thus introduce the following dimensionless notation:

$$P=\frac{p}{p_0},\quad y=\omega_0 t,\quad \sigma=\frac{\beta p_0\omega_0}{\rho_0 c_0^3}x=\frac{x}{\bar{x}}, \tag{3}$$

where P is the sound pressure normalized by a characteristic pressure amplitude p_0, y is the dimensionless time normalized by a characteristic frequency ω_0, and σ is the dimensionless distance normalized by $\bar{x}$, which is the lossless plane-wave shock formation distance for a sinusoidal source condition. In terms of this notation, Eq. (1) becomes

$$\frac{\partial P}{\partial \sigma}=NP\frac{\partial P}{\partial y}+A\frac{\partial^2 P}{\partial y^2}, \tag{4}$$

$$A=\frac{\delta\omega_0\rho_0}{2\beta p_0}, \quad N=1. \tag{5}$$

The linear term on the left side of Eq. (4) (P_σ) is a linear wave propagation term. The first term on the right side of the equation (NPP_y) is a nonlinear term that accounts for quadratic nonlinearity. The coefficient N in front of the nonlinear term has unity value and it is added in the equation to tag the nonlinear term so that the latter can be identified later in the paper. The second term on the right side of the equation (AP_{yy}) is a linear dissipation term that accounts for thermoviscous absorption. It should be mentioned that additional linear terms can be added to Eq. (4) to account for other phenomena such as geometrical spreading and attenuation due to relaxation.[2,3]

If both coefficients N and A are zero, the solution of Eq. (4) represents a wave that maintains its shape throughout the propagation. If $N=0$, the BE reduces to a diffusion-like equation (σ and y exchange the roles they have in the usual heat conduction equation); the solution represents a wave that diffuses as it propagates. Finally, for $A=0$ Eq. (4) is a first-order nonlinear progressive wave equation; the wave distorts as it propagates and eventually shocks are formed.[3]

BE has an exact analytical solution known as the Cole–Hopf solution[4,5,3] In addition to the Cole–Hopf solution, a great number of alternative methods and numerical algorithms exist in the literature.[6–9] A detailed survey of the existing computational methods can be found in Ref. 3. In all these cases, the pressure at the source ($\sigma=0$) is prescribed and the solution represents the pressure waveform at any point ($\sigma>0$) along the propagation path.

In many cases, however, the time waveform at $\sigma=0$ is not known. Consider, for example, a nondeterminstic source condition, i.e., a stochastic process. A process is characterized by its power spectral density (PSD); the time waveform at the source is not known. One can also consider cases of deterministic signals, where the PSD of the signal is known but not the signal itself. In aircraft-noise measurements, for example, the measured signal itself is not reported; the PSD is given instead.

A substantial complication arises when the source condition is defined by a PSD instead of a time waveform. A given time waveform has a unique PSD, but a given PSD does not correspond to a unique time waveform. A unique source condition in the time domain cannot be found from the PSD because the phase information is missing. In general, there are two ways of tackling the problem of the missing phase information. The first is to assign somehow a phase content to the signal and then employ the Cole–Hopf solution or any of the existing methods to solve the BE. In this case, the solution gives the time waveform along the propagation path. The second approach is to predict the evolution of the PSD directly by solving an appropriate form of the BE. In this case, the solution gives the evolution of the PDS along the propagation path.

Pestorius *et al.*[10] followed the first approach. Assuming that the signal had a uniformly random phase distribution, they reconstructed several time waveforms from an initial PSD. Subsequently, the waves were propagated numerically. The PSDs predicted from the different waveforms after the propagation were substantially different. This numerical experiment shows that the phase content of the waveform is essential for the prediction of its evolution. Therefore, if the first approach is to be followed, one must solve the BE a great many times with different time waveforms at the source, all having the same PSD, and then average the resulting PSDs. This approach is very time consuming.

The vast majority of researchers has focused on the second approach, i.e., working directly with the PSD or its Fourier transform counterpart, the autocorrelation function (AF). Rudenko and Chirkin[11] have contributed a closed analytical solution for the evolution of the AF (or the PSD) of a normal stochastic process in the shock-free region, and in the absence of absorption ($A=0$). Their solution is a landmark and the results from the present work are compared with it in Sec. VI. Later, Rudenko[12] presented a method to include thermoviscous attenuation. An auxiliary AF is defined and, for short distances or smooth time profiles, its evolution is approximated by the solution of a linearized diffusion equation. The AF is found at the next step via a nonlinear transformation of the auxiliary AF. The work of Crighton and Bashforth,[13] who presented a series solution, should also be mentioned. They used a truncated Taylor series for the PSD as the starting point of a prediction scheme. Starting with a truncated series, however, imposes severe limitations in the range of propagation distance.

Many methods that predict the evolution of the AF directly are based on so-called closure hypotheses. It is known that the nonlinear equation in the time domain (in this case the Burgers equation) can be transformed into an infinite number of nested linear equations in which the unknowns are the AF (or the PSD) and higher order moments (or higher order cross spectra).[14] The evolution of the AF depends on a third order moment, which in turn depends on fourth order moments and so on. Therefore, the equation for AF remains ''unclosed.'' A similar problem has been reported earlier in the theory of hydrodynamic turbulence. One of the many models for turbulence is based on the BE[15] [in the original form of the BE, σ and y exchange the role they have in Eq. (4)]. (A discussion of the history of the BE can be found in Ref. 3, pp. 118–126.) The analog of the unclosed equation of AF (or PSD) is the Karman–Howarth equation in the theory of hydrodynamic turbulence. A considerable amount of work has been published on various closure schemes. Researchers[16–21] in the field of turbulence have replaced the unclosed equation for AF (or PSD) by a model equation, expressing the unknown third-order moment (or cross spectra) with known quantities that depend on the AF. Perhaps the most well known are Kraichnan's ''direct interaction'' approximation and the method of ''Markov stochastic coupled models.'' It has been reported[14] that in these models the AF ceases to be analytic after a finite time as a result of the discontinuities that are created. Although these methods have not been used in acoustical applications, their results indicate that similar closure schemes would lead to AF that ceases to be analytic after a finite propagation distance as a result of shock formation. The only closure scheme that has been used in real acoustical applications, to the best of the

authors' knowledge, is the so-called quasinormal hypothesis. A Gaussian signal at the source becomes non-Gaussian as it propagates because of nonlinear effects. According to quasinormal hypothesis, the signal remains almost Gaussian throughout the propagation. Morfey and Howell[22,23] have used this hypothesis in acoustical applications. Comparison with aircraft-noise measurements did not provide a very good agreement. The hypothesis has a serious flaw; one predicts negative PSD, which is by definition a positive quantity. Jeng[24] and Ogura[25] have also observed that artifact.

In the present work a new formulation of the unclosed system of equations is presented and a method to extract results from a subset of the equation is outlined. The novel element of the work presented here is that it does not require any closure hypothesis. The evolution of the AF is predicted under the combined effects of nonlinearity and thermoviscous attenuation. The method can be further extended to include geometrical spreading. The method presented here requires knowledge of the AF and the higher order moments at the source. If the source condition has Gaussian characteristics (it is a Gaussian noise or a Gaussian stochastic process), then all moments can be computed from the AF; thus, only knowledge of the AF is required. The method, however, is not restricted to source conditions with Gaussian characteristics. It is applicable to all cases where the appropriate moments at the source are known. The method is outlined in Secs. II and III and its numerical implementation is described in Sec. IV. In Sec. VI numerical results from the method and their agreement with known solutions are presented in the preshock region. The restriction to the preshock region is attributed to computational limitations and not to the method, which is in principle applicable in the entire range of propagation.

II. A NEW FORMULATION OF THE INFINITE SET OF LINEAR EQUATIONS

It is known[14] that the BE [Eq. (4)] can be transformed into an infinite number of nested linear equations. The transformation is shown in subsection A. A new formulation for the system of equations is presented in subsection B. The new formulation reveals the structure of the system and allows us to represent the infinite set of equations by a single recursion equation.

A. The infinite set of equations

Consider the autocorrelation function for $P(\sigma,y)$

$$R_{12}(\psi)=\langle P_1P_2\rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P_1P_2W_{P_1P_2}(P_1,P_2;\sigma,\psi)dP_1\,dP_2, \tag{6}$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_0^T P_1P_2\,dy, \tag{7}$$

where $P_1=P(\sigma,y)$ is the pressure at time y, $P_2=P(\sigma,y+\psi)$ is pressure P_1 delayed by time ψ, and $W_{P_1P_2}(P_1,P_2;\sigma,\psi)$ is the second-order joint probability density function at ψ.[26] In the present paper it is assumed that P is either a stationary and ergodic stochastic process or a deterministic signal. In the former case the ensemble average represented in Eq. (6) can be computed by Eq. (7). In the latter case the averaging is directly computed by Eq. (7).

A statistical version of the BE is obtained in terms of the autocorrelation function by averaging Eq. (4)

$$\frac{\partial}{\partial\sigma}R_{12}(\psi)=N\frac{\partial}{\partial\psi}[-R_{122}(-\psi)+R_{122}(\psi)] + A\frac{\partial^2}{\partial\psi^2}[R_{12}(-\psi)+R_{12}(\psi)], \tag{8}$$

where $R_{122}(\psi)=\langle P_1,P_2P_2\rangle$ is a third-order joint moment. The evolution of R_{12} cannot be computed by Eq. (8), because the third-order joint moment R_{122} is unknown. The evolution of the unknown third-order moment can be described by a similar averaging of the BE that includes the terms $N(\partial/\partial\psi)R_{1122}(-\psi)$ and $N(\partial/\partial\psi)R_{1222}(\psi)$, where $R_{1122}=\langle P_1P_1,P_2P_2\rangle$ and $R_{1222}=\langle P_1,P_2P_2P_2\rangle$ are joint moments of fourth order. Similarly, the evolution of fourth-order moments requires knowledge of fifth-order moments and so on. Therefore, the system of the coupled equations remains unclosed.

B. A new formulation

Previous researchers[16–23] have employed a hypothesis to close the set of the nested equations. In the present work we continue to derive evolution equations for the ever higher order joint moments. The rather tedious calculations reveal a certain structure in the set of equations.

Let us first introduce the notation $E_{m,n}$ for the joint moments

$$E_{m,n}(\psi)=\langle P_1^mP_2^n\rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P_1^mP_2^nW_{P_1P_2}(P_1,P_2;\sigma,\psi)dP_1\,dP_2 = \lim_{T\to\infty}\frac{1}{T}\int_0^T P^m(y)P^n(y+\psi)dy. \tag{9}$$

According to the new notation $R_{12}=E_{1,1}$, $R_{122}=E_{1,2}$ and so forth. The order $\overline{N}$ of the joint moment is equal to the sum of the indices, i.e., $\overline{N}=m+n$. There are $\overline{N}-1$ moments of order $\overline{N}$. Finally, the joint moments $E_{m,n}$ possess the following symmetry property:

$$E_{m,n}(\psi)=E_{n,m}(-\psi). \tag{10}$$

Using the new notation, we continue to derive evolution equations [such as Eq. (8)] for the ever higher order joint moments. The calculations result in a recursion equation. This equation shows the dependency of the joint moment $E_{m,n}$ on the higher order moments and thus reveals the structure of the set of equations. More specifically, consider Eq. (4) at two instances y and $y+\psi$, respectively

$$\frac{\partial P(\sigma,y)}{\partial\sigma}=NP(\sigma,y)\frac{\partial P(\sigma,y)}{\partial y}+A\frac{\partial^2P(\sigma,y)}{\partial y^2}, \tag{11}$$

$$\frac{\partial P(\sigma,y+\psi)}{\partial \sigma}=NP(\sigma,y+\psi)\frac{\partial P(\sigma,y+\psi)}{\partial y}+A\frac{\partial^2 P(\sigma,y+\psi)}{\partial y^2}. \tag{12}$$

Multiply the first equation with $mP^{m-1}(y)P^n(y+\psi)$, the second with $nP^{n-1}(y+\psi)P^m(y)$, and add the two together. The statistical averaging of the resulting equation gives

$$\left\langle mP^{m-1}(y)P^n(y+\psi)\frac{\partial P(\sigma,y)}{\partial \sigma}+nP^m(y)P^{n-1}(y+\psi)\frac{\partial P(\sigma,y+\psi)}{\partial \sigma}\right\rangle$$
$$=\left\langle NmP^{m-1}(y)P^n(y+\psi)P(\sigma,y)\frac{\partial P(\sigma,y)}{\partial y}+NnP^m(y)P^{n-1}(y+\psi)P(\sigma,y+\psi)\frac{\partial P(\sigma,y+\psi)}{\partial y}\right\rangle$$
$$+\left\langle AmP^{m-1}(y)P^n(y+\psi)\frac{\partial^2 P(\sigma,y)}{\partial y^2}+AnP^m(y)P^{n-1}(y+\psi)\frac{\partial^2 P(\sigma,y+\psi)}{\partial y^2}\right\rangle. \tag{13}$$

The order of the terms in the above equation should be noted. The term on the left-hand side and the second term on the right-hand side are of order $m+n$, while the first term on the right-hand side, which is tagged with the coefficient N and originates from the nonlinear term in the BE, is of order $m+n+1$. Use of Eq. (9) yields

$$\frac{\partial}{\partial \sigma}E_{m,n}(\sigma,\psi)=N\frac{\partial}{\partial \psi}\left[-\frac{m}{m+1}E_{n,m+1}(\sigma,-\psi)+\frac{n}{n+1}E_{m,n+1}(\sigma,\psi)\right]+A\frac{\partial^2}{\partial \psi^2}[E_{n,m}(\sigma,-\psi)+E_{m,n}(\sigma,\psi)]$$
$$-mA\left\langle\frac{\partial P^{m-1}(\sigma,y)}{\partial y}\frac{\partial P(y)}{\partial y}P^n(y+\psi)\right\rangle-nA\left\langle P^m(y)\frac{\partial P(y+\psi)}{\partial y}\frac{\partial P^{n-1}(y+\psi)}{\partial y}\right\rangle. \tag{14}$$

Equation (14) couples two moments of order $m+n$ ($E_{m,n}$ and $E_{n,m}$) and two moments of order $m+n+1$ ($E_{m,n+1}$ and $E_{n,m+1}$). The last two terms of the right-hand side of Eq. (14), which are not expressed in terms of joint moments $E_{m,n}$, originate from the linear dissipation term, are of order $m+n$ and vanish for ($m=1$, $n=1,2,...$) and ($n=1$, $m=1,2,...$), respectively. These two terms have been neglected in this work for reasons that relate to the sequence of computations in the presented method as outlined in Sec. III combined with the fact that they vanish for ($m=1$, $n=1,2,...$) and ($n=1$, $m=1,2,...$). However, the relative importance of the neglected terms compared to the other terms justifies a thorough study, particularly if one desires to use the equations derived in this subsection in connection with a method different than the one outlined in the present work. The symmetry property of $E_{m,n}$ [Eq. (10)] is used to eliminate $E_{n,m}(\sigma,-\psi)$ and $E_{n,m+1}(\sigma,-\psi)$ from Eq. (14) and to obtain for any $m\geqslant 1$ and $n\geqslant 1$ the final equation

$$\frac{\partial}{\partial \sigma}E_{m,n}(\sigma,\psi)=N\frac{\partial}{\partial \psi}\left[-\frac{m}{m+1}E_{m+1,n}(\sigma,\psi)+\frac{n}{n+1}E_{m,n+1}(\sigma,\psi)\right]+2A\frac{\partial^2}{\partial \psi^2}E_{m,n}(\sigma,\psi). \tag{15}$$

This recursion equation is derived for the first time. Because it results from the statistical averaging of the BE, it will be referred to as the *statistical Burgers equation* (SBE). It is the representative equation for the infinite set of equations and reveals that any moment $E_{m,n}$ depends on the second derivative of $E_{m,n}$ and on first derivatives of the moments $E_{m+1,n}$ and $E_{m,n+1}$. Please note that although the main intent is to apply the SBE to propagation of stochastic processes, Eq. (15) also applies to deterministic signals. $E_{m,n}$ represent moments of stationary and ergodic stochastic processes or of deterministic signals.

As opposed to the BE [Eq. (4)], the differential operators

TABLE I. Graphical illustration of the correspondence between terms in the BE [Eq. (4)] and SBE [Eq. (15)].

BE:	$\frac{\partial P}{\partial \sigma}$	=	$NP\frac{\partial P}{\partial y}$	+	$A\frac{\partial^2 P}{\partial y^2}$
SBE:	$\frac{\partial}{\partial \sigma}E_{m,n}$	=	$N\frac{\partial}{\partial \psi}\left[-\frac{m}{m+1}E_{m+1,n}+\frac{n}{n+1}E_{m,n+1}\right]$	+	$2A\frac{\partial^2}{\partial \psi^2}E_{m,n}$
			Couples higher order moments Accounts for nonlinear distortion		Accounts for thermoviscous absorption

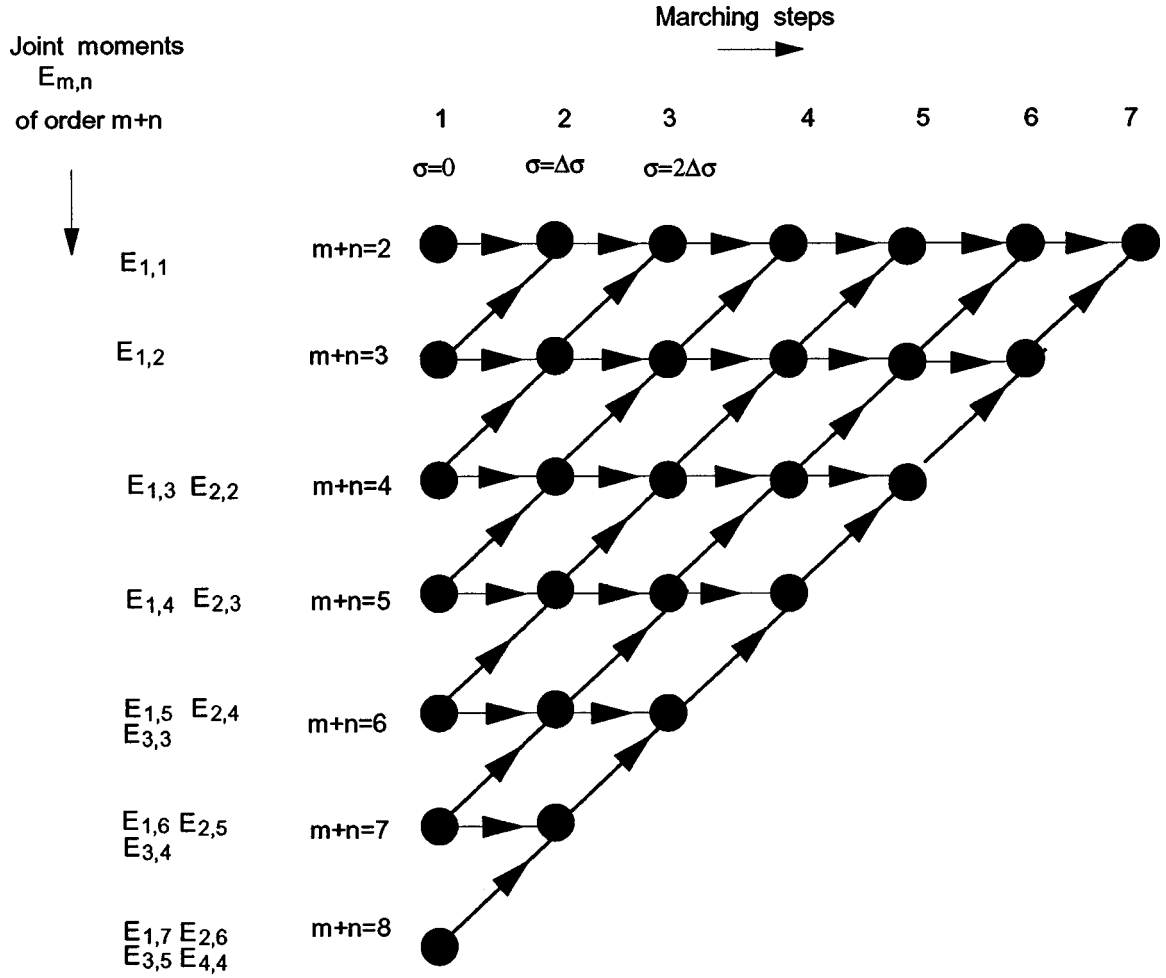


FIG. 1. Graphical illustration of the new method showing the interactions of the joint moments $E_{m,n}$. The interaction paths form a triangle with the joint moments on the vertical side and the marching solution steps on the horizontal. The arrows indicate the transfer of information between the joint moments. If all the joint moments up to order $m+n=\overline{N}$ are known at the source ($\sigma=0$), the autocorrelation function $E_{1,1}$ can be computed at each propagation step $j=1,...,\overline{N}-1$.

in Eq. (15) are all linear. Comparing the two, one can see that the first term on the right side of Eq. (15) comes from the nonlinear term of the BE, the second term on the right comes from the linear dissipation term of the BE, and the term on the left side of Eq. (15) comes from the linear propagation term in the BE (see Table I). The joint moments in the term that corresponds to the nonlinear BE term are of higher order than the joint moment on the left side of Eq. (15). The nonlinear effects are responsible for the coupling of the equations, which results in the unclosed system. Thus, although its differential operator is linear, the first term on the right side of the SBE accounts for the nonlinear distortion. On the other hand, the joint moments in the term that corresponds to the linear dissipation term of BE are of the same order as the one on the left side of Eq. (15). The linear dissipation term of the BE does not contribute to the coupling of the joint moments. As a result, in the absence of nonlinearity ($N=0$) the system of equations represented by the SBE reduces to a single equation for $E_{1,1}$, which is identical in form with the BE for $N=0$.

III. RESULTS FROM A SUBSET OF THE EQUATIONS

The main advantage of the new formulation of the infinite system of equations is that it reveals their structure. The insight provided by the SBE allows us to use a subset of the equations to extract numerical results without making any approximation about the equations outside the subset. It should be emphasized that all the previous researchers have closed the system of equations by making severe approximations about the physics of the problem.

We want to develop a marching algorithm to calculate the joint moments $E_{m,n}$ as σ increases. Consider a first-order accurate Taylor series expansion of $E_{m,n}(\sigma,\psi)$ around $\sigma=\sigma_0+\Delta\sigma$

$$E_{m,n}|_{\sigma_0+\Delta\sigma}=E_{m,n}|_{\sigma_0}+\frac{\partial}{\partial\sigma}E_{m,n}|_{\sigma_0}\Delta\sigma+O[(\Delta\sigma)^2]. \tag{16}$$

For the first-order spatial derivative in the expansion substitute the right side of the SBE [Eq. (15)] evaluated at σ_0

$$E_{m,n}|_{\sigma_0+\Delta\sigma}=E_{m,n}|_{\sigma_0}+\Delta\sigma N\frac{\partial}{\partial\psi}\left[-\frac{m}{m+1}E_{m+1,n}+\frac{n}{n+1}E_{m,n+1}\right]|_{\sigma_0}+2A\frac{\partial^2}{\partial\psi^2}E_{m,n}|_{\sigma_0}+O[(\Delta\sigma)^2]. \tag{17}$$

The above equation reveals that if all moments up to order $m+n+1$ are known at σ_0, then all the moments of lower order can be computed at $\sigma_0+\Delta\sigma$. Alternatively, all moments up to order $m+n+1$ have to be known at σ_0 in order to calculate the moments up to order $m+n$ at $\sigma_0+\Delta\sigma$.

We are ultimately interested in the evolution of just one of the joint moments, namely of $E_{1,1}$, the autocorrelation function. If $E_{1,1}$ is to be computed at σ, then $E_{1,1}$, $E_{1,2}$, and $E_{2,1}$, i.e., all the second- and third-order moments are needed at point $\sigma-\Delta\sigma$. And, for the latter to be known, all the moments up to order four are needed at point $\sigma-2\Delta\sigma$, and so forth. Continuing to march backward, one finally reaches the source. If all the joint moments up to order $\overline{N}$ are known at the source ($\sigma=0$), the autocorrelation function can be computed at each propagation step $\sigma=0,\Delta\sigma,2\Delta\sigma,\ldots(N-2)\Delta\sigma$. The idea is illustrated graphically in Fig. 1.

The joint moments at the source can be computed (i) if the time history of the source signal is given; (ii) if the source signal is Gaussian noise with known AF; or (iii) if the source condition is a Gaussian stationary and ergodic stochastic process with known AF. The details are given in the Appendix.

The method described here allows us to extract results by solving numerically a selected subset of the infinite set of equations represented by the statistical Burgers equation [Eq. (15)] without providing any closure hypothesis.

IV. NUMERICAL IMPLEMENTATION

For the numerical implementation of the method presented in this paper, the standard Crank–Nicolson scheme[27] is employed. The SBE can be discretized as follows:

$$E^{j+1}_{m,n,i}=E^{j}_{m,n,i}+\frac{d}{2}[E^{j}_{m,n,i+1}-2E^{j}_{m,n,i}+E^{j}_{m,n,i-1}]+\frac{d}{2}[E^{j+1}_{m,n,i+1}-2E^{j+1}_{m,n,i}+E^{j+1}_{m,n,i-1}]+S. \quad (18)$$

The superscript j indicates the marching step, and the subscript i indicates the grid point along the ψ axis. The coefficient d is the diffusion number ($d=2A\Delta\sigma/\Delta\psi^2$) and S is the following source term:

$$S=-\frac{m}{(m+1)}\frac{\tilde{c}}{2}[\overline{E}_{m+1,n,i+1}-\overline{E}_{m+1,n,i-1}]+\frac{n}{(n+1)}\frac{\tilde{c}}{2}[\overline{E}_{m,n+1,i+1}-\overline{E}_{m,n+1,i-1}], \quad (19)$$

where $\tilde{c}$ is the Courant number[27] ($\tilde{c}=N\Delta\sigma/\Delta\psi$) and the overbars indicate that the joint moments included in the source term will be evaluated either at the present propagation step $j+1$ ($E^{j+1}_{m+1,n}$,$E^{j+1}_{m,n+1}$) or at the previous propagation step j ($E^{j}_{m+1,n}$,$E^{j}_{m,n+1}$). The higher order moments $\overline{E}_{m+1,n}$ and $\overline{E}_{m,n+1}$ in the source term will always be known when the joint moment $E_{m,n}$ in Eq. (18) is computed. They will always be known at the previous propagation step j, and depending on the sequence of the calculations they will also be known at the present propagation step $j+1$.

Based on properties of the joint moments (see the Appendix) the value of $E_{m,n}$ at the boundaries of the computational domain ($\psi=0$ and $\psi=\Psi$) is specified by the following boundary conditions:

$$E_{m,n}(\psi)=0 \quad \text{at } \psi=0,\Psi \quad \text{if } m+n \text{ is odd}$$
$$\frac{\partial E_{m,n}}{\partial\psi}=0 \quad \text{at } \psi=0,\Psi \quad \text{if } m+n \text{ is even.} \quad (20)$$

The numerical scheme employed in Eq. (18) is second-order accurate in both time and space, and thus the numerical error is of the order of $O[(\Delta\psi)^2]$ and $O[(\Delta\sigma)^2]$. The numerical scheme results in a tridiagonal system of algebraic equations at each marching step. The resulting system is solved by using the Thomas algorithm.[27]

The Crank–Nicolson scheme is unconditionally stable. However, oscillations can occur in the computation of the source term in Eq. (18) and affect the stability of the algorithm. In cases of weak nonlinearity, the relative importance of the terms in Eq. (18) containing the diffusion number d is greater than that of the source term, and therefore the oscillations are damped by the dissipation mechanism. In cases of strong nonlinearity, however, caution should be exercised. In order to avoid contamination of the results by numerical oscillations in the source term, the latter is averaged before it is included in Eq. (18). Although the method outlined in Secs. II and III is in principle applicable for the entire range of propagation, the computational limitation imposed by the oscillations in the source term restricts the numerical results presented here in the preshock region for cases of strong nonlinearity. The numerical results obtained by the presented algorithm are shown in Sec. VI, as well as their agreement with known analytical solutions.

V. THE EFFECTS OF NONLINEARITY AND DISSIPATION ON THE EVOLUTION OF THE AUTOCORRELATION FUNCTION FOR A SINUSOIDAL SOURCE CONDITION

The effects of nonlinearity and dissipation are usually demonstrated in the time domain. The purpose of this section is to demonstrate their effects in the autocorrelation domain. To do that we have chosen the well-studied case of a pressure signal that is sinusoidal at the source. The evolution of the pressure time waveform, as distance increases, is computed by known analytical solutions. The corresponding evolution of the autocorrelation function is computed by averaging these analytical time-domain solutions. In this section it is shown how nonlinearity and thermoviscous attenuation manifest themselves in both the time and autocorrelation domain. Results from the new method are shown in the next section.

The evolution of a sinusoidal source signal [$p(0,t')=p_0\sin\omega_0 t'$] is considered for the following cases:

(1) Small-signal waves in a dissipative fluid [$N=0$, $A\neq 0$ in Eq. (4)]; results are shown in Figs. 2(a) and (b);
(2) Finite-amplitude waves in a dissipative fluid [$N\neq 0$, $A\neq 0$ in Eq. (4)]; results are shown in Figs. 2(c) and (d); and

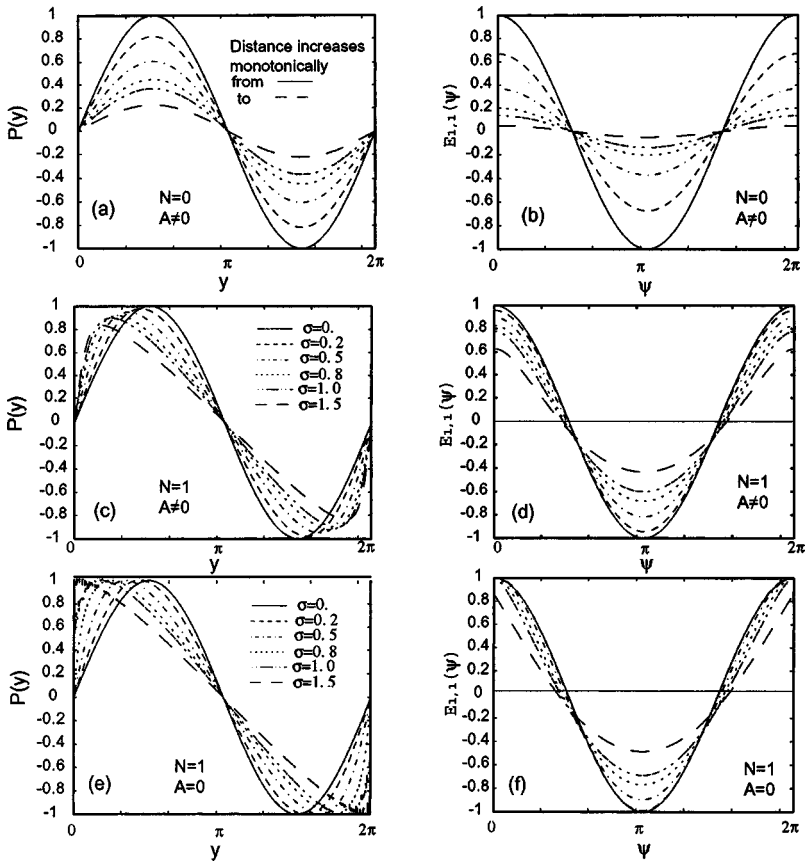


FIG. 2. The effects of nonlinearity and/or thermoviscous attenuation on the pressure waveform $P(y)$ and on the autocorrelation function $E_{1,1}(\psi)$ for a sinusoidal source signal at various propagation distances σ from the source. Time waveforms are shown on the left, autocorrelation functions on the right. (a), (b): Small-signal wave in a dissipative fluid ($N=0$, $A\neq 0$). (c), (d): Finite-amplitude wave in a dissipative fluid ($N=1$, $A\neq 0$). (e), (f): Finite-amplitude wave in a fluid with no dissipation except at shocks ($N=1$, $A=0$).

(3) Finite-amplitude waves in a fluid with no dissipation except at shocks [$N\neq 0$, $A=0$ in Eq. (4)]; results are shown in Figs. 2(e) and (f).

In all three cases closed analytical solutions exist in the time domain, and the results are presented on the left column of Fig. 2. The evolution of the AF is computed by averaging those analytical solutions, and the results are presented on the right column of Fig. 2.

All time-domain results $[P(\sigma,y)]$ are normalized by the source-amplitude pressure p_0 and have unit amplitude at the source. All autocorrelation-domain results $[E_{1,1}(\sigma,\psi)]$ are normalized to have unit mean-square value at the source $[E_{1,1}(0,0)=1]$. Time quantities are normalized by the frequency ω_0 of the periodic wave and distance is normalized, for the finite-amplitude cases, by the shock formation distance, $\bar{x}$.

Propagation of small-signal waves in a dissipative fluid ($N=0$, $A=1$) is considered first. The sinusoidal wave retains its shape and decays by absorption as it propagates [Fig. 2(a)]. The same behavior is observed in the autocorrelation domain. The function $E_{1,1}$ retains its original cosine shape and decays by absorption as it propagates [Fig. 2(b)]. The following observations should be noted: (i) The maximum and the minimum value in the autocorrelation function have the same absolute value at each propagation distance. (ii) The zero crossings do not change position with distance. (iii) The value of $E_{1,1}$ at $\psi=0$ decreases with distance because of the absorption.

Propagation of small-signal waves in a dissipative fluid ($N=1$, $A=0.1$) is considered next. We have found the time waveforms for $A=0.1$ by evaluating the closed analytical solution of the BE known as the Cole–Hopf solution.[4,5] The steepening of the time waveform can be clearly observed in the time domain [Fig. 2(c)]. The effects of nonlinearity on the autocorrelation function can also be observed [Fig. 2(d)]. The initial cosine shape of the autocorrelation changes. The maximum and the minimum value do not have the same absolute value; the minimum value decreases faster than the maximum value. The zero crossings move toward the ends of the time ψ interval, as distance increases. The value of $E_{1,1}$ at $\psi=0$ decreases because of the presence of absorption. Absorption is responsible for the uniform decrease of the autocorrelation function, while nonlinear effects cause its distortion.

Finally, propagation of finite-amplitude wave in a fluid with no dissipation except at shocks is considered. The distortion of the time wave can be described by the Poisson solution before the formation of shocks ($\sigma\leqslant 1$) and by weak shock theory afterwards. For a sinusoidal source condition the solution can be expressed in the time domain in a closed analytical form[28] as follows:

$$P(\sigma,y)=\sum_{n=1}^{\infty} B_n \sin ny, \tag{21}$$

where the coefficients B_n reduce to the Fubini coefficients[28] for $\sigma<1$. The evolution of the autocorrelation function may be expressed as

$$E_{1,1}(\sigma,\psi)=\sum_{n=1}^{\infty} B_n \cos n\psi. \tag{22}$$

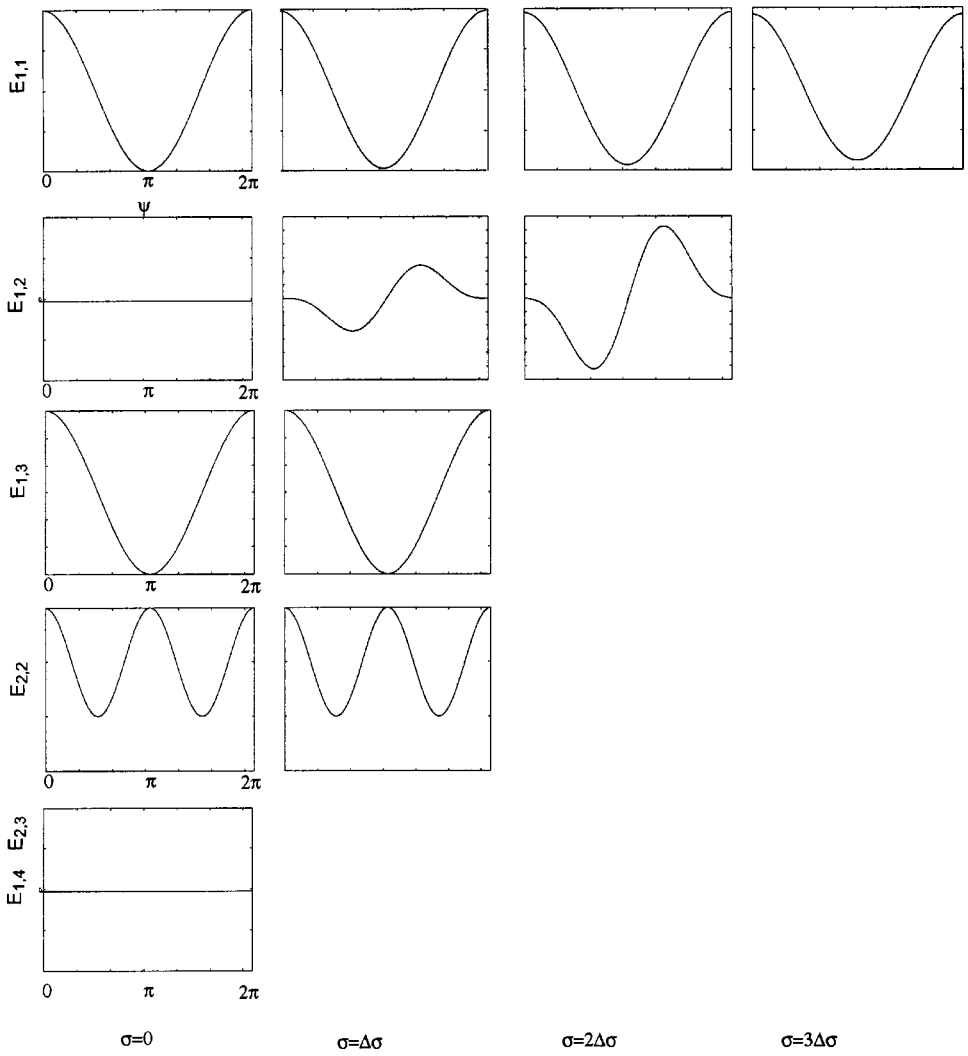


FIG. 3. The evolution of the joint moments $[E_{m,n}(\psi)]$ for a sinusoidal source signal $[P(y)]$ at various propagation distances σ from the source ($\Delta\sigma$ is the propagation step.)

The distortion of the autocorrelation function can be observed in Fig. 2(f) and it is similar to that of the previous case. It should be noted that $E_{1,1}(0)$ remains unity for $\sigma < 1$, because of the absence of absorption, and that is decreases for $\sigma \geqslant 1$ as a result of losses at the shock fronts.

The comparison between the time-domain evolution and the autocorrelation domain evolution reveals that absorption and nonlinear effects can be observed in the autocorrelation domain as well. Absorption is responsible for a uniform decrease of the autocorrelation function, while nonlinear effects are responsible for its distortion. Furthermore, working in the autocorrelation domain has some advantages. First, the autocorrelation function is always a continuous function, while the time waveform is discontinuous at the shock fronts. As a result, the time-domain series in Eq. (21) needs considerably more terms for its numerical convergence than the series in Eq. (22) because of the development of the steep shock front. Second, the rms value of the signal, and thus its intensity, are easily computed by the value of $E_{1,1}$ at $\psi = 0$.

VI. RESULTS

In the following, numerical results from the method described in Secs. II, III, and IV are presented. The results are compared in the preshock region with benchmark analytical solutions for the following source conditions: (i) a sinusoidal signal and (ii) a Gaussian stochastic process with known PSD at the source.

All the results presented in this section are expressed in dimensionless variables. The characteristic pressure is chosen so that $E_{1,1}(0,0) = 1$, and the characteristic angular frequency ω_0 is either the angular frequency of the sinusoidal source signal or the inverse correlation time of the process' autocorrelation function at the source.

A. Sinusoidal source signal

For a sinusoidal source signal all the joint moments at the source can be computed as outlined in the Appendix. The joint moments of odd order are identically zero at the source, while the even-order joint moments are even functions at the source. The joint moments evolve according to the SBE. As a result, even-order joint moments remain even functions throughout the propagation, while odd-order moments (zero at the source) become and remain odd functions throughout the propagation.

Figure 3 shows the evolution of higher order moments and how they contribute to the evolution of the AF. (Figure 3 does not show the actual numerical results. Very large propagation steps—$\Delta\sigma = 0.1$ compared to $\Delta\sigma = 0.02$ typically used in the numerical scheme—have been used in order to

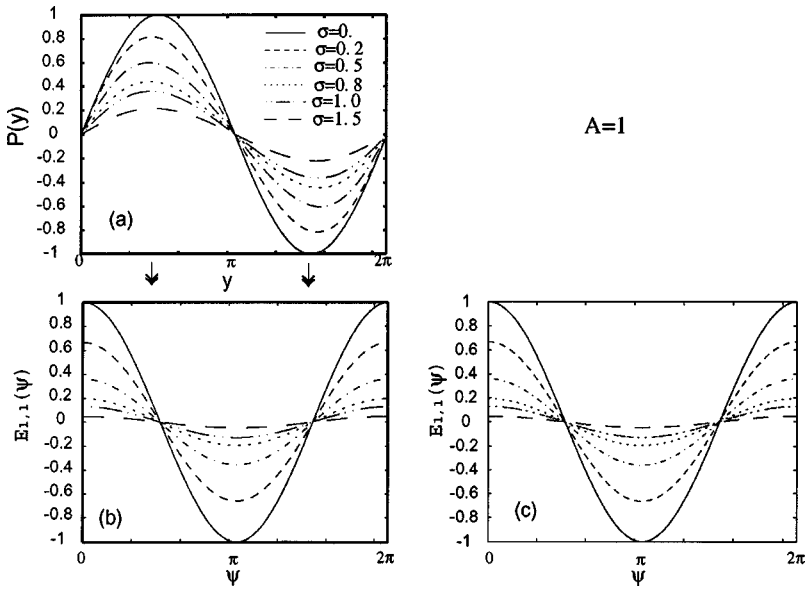


FIG. 4. Comparison of numerical results with known solutions for the case of a sinusoidal source signal $P(y)=\sin y$ and for $A=1$. The evolution of the pressure waveform at various propagation distances σ is shown in (a). The corresponding $E_{1,1}(\psi)$ shown in (b) has been obtained by averaging the time-domain results shown in (a). The results shown in (c) have been obtained directly in the joint moments domain by the new method.

demonstrate noticeable differences with just a few joint moments at the source.) A final remark for the location of the zero crossings and the extrema of the joint moments should be added. The number of extrema as well as the number of zero crossings of a joint moment $E_{m,n}$ remains the same throughout the propagation. In the evolution of a joint moment of even order the locations of the extrema remain the same, while the zero crossings locations change. For the odd-order joint moments, the reverse is true.

Figure 4 shows the evolution of the time waveform $P(y)$ and of the autocorrelation function $E_{1,1}(\psi)$ for a weak wave ($A=1$) launched as a pure sinusoid. Figure 4(a) shows how the pressure waveform $P(y)$ changes with propagation distance σ. For the time-domain calculations we have used the numerical algorithm of Ref. 9. Figure 4(b) shows the corresponding $E_{1,1}(\psi)$, as obtained by averaging the time waveforms according to Eq. (9). Figure 4(c) shows the same $E_{1,1}(\psi)$, but as obtained by the new method. According to the new method the evolution of the AF is computed directly in the autocorrelation domain knowing all joint moments at the source, instead of the source signal. Note that the joint moments at the source can be computed from the AF at the source using Eq. (A5). It can be observed that the results for the AF [shown in (b)] obtained from the time-domain calculations are in good agreement with the numerical results for the AF obtained by the new method [shown in (c)].

Comparisons for smaller values of A (i.e., for stronger waves) are shown in Figs. 5 ($A=0.1$), 6 ($A=0.01$), and 7 ($A=0.001$). It can be observed that the numerical results are in reasonably good agreement with the time-domain calculations.

The computation were performed on Sparc 2000 SUN stations. The computation time increases noticeably with propagation distance but only slightly as A decreases. The computation time for the results presented here was approximately 7 min. The propagation step used in the computations shown in Figs. 5–7 was $\Delta\sigma=0.02$ and joint moments up to order $\bar{N}=41$ were computed at the source. For the computations shown in Fig. 4 (case of weak nonlinearity) the propa-

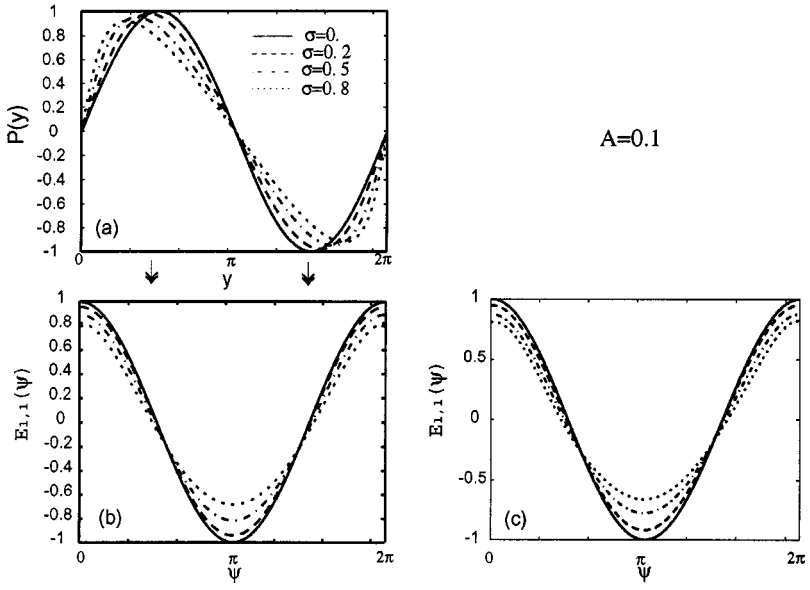


FIG. 5. Comparison of numerical results with known solutions for the case of a sinusoidal source signal $P(y)=\sin y$ and for $A=0.1$. The evolution of the pressure waveform at various propagation distances σ is shown in (a). The corresponding $E_{1,1}(\psi)$ shown in (b) has been obtained by averaging the time-domain results shown in (a). The results shown in (c) have been obtained directly in the joint moments domain by the new method.

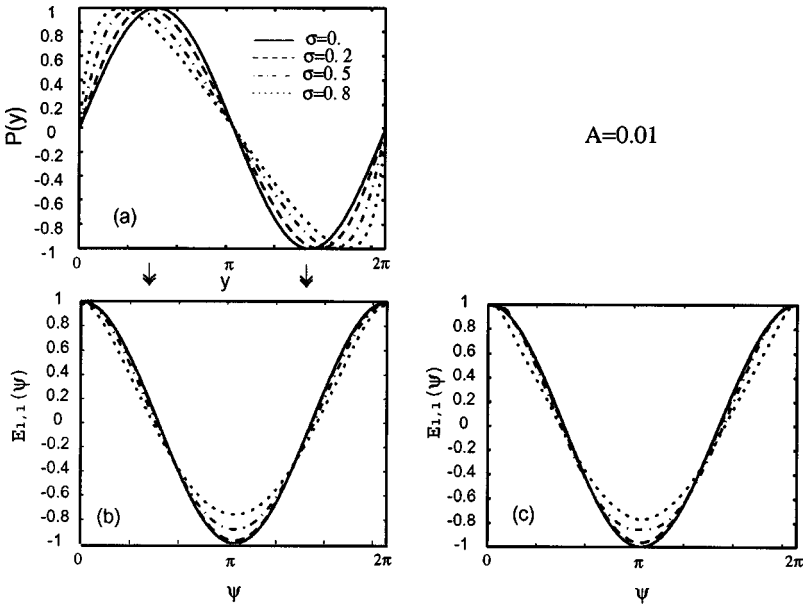


FIG. 6. Comparison of numerical results with known solutions for the case of a sinusoidal source signal $P(y)=\sin y$ and for $A=0.01$. The evolution of the pressure waveform at various propagation distances σ is shown in (a). The corresponding $E_{1,1}(\psi)$ shown in (b) has been obtained by averaging the time-domain results shown in (a). The results shown in (c) have been obtained directly in the joint moments domain by the new method.

gation step was $\Delta\sigma=0.05$ and the required joint moments at the source up to order $\overline{N}=31$. Finally, it should be added that for each result presented here, many simulations have been performed to assure that convergence with regards to $\Delta\sigma$ and $\overline{N}$ has been achieved.

The effects of nonlinearity and dissipation can be observed in the autocorrelation domain as explained in Sec. V. (i) The value of $E_{1,1}(\psi)$ at $\psi=0$, which indicates the dissipation losses, decreases noticeably with propagation distance for weak nonlinearity ($A=1$ shown in Fig. 4 and $A=0.1$ shown in Fig. 5), while it remains almost unity for strong nonlinearity ($A=0.01$ shown in Fig. 6 and $A=0.001$ shown in Fig. 7). (ii) The original cosine shape of the AF curve changes little, apart from decaying in amplitude, for weak nonlinearity ($A=1$) but distorts noticeably in the case of strong nonlinearity. In the former case the minimum value of the curve has almost the same absolute value as its maximum at each propagation step. For strong nonlinearity the absolute value of the curve's minimum decreases with propagation distance faster than the value of its maximum.

Finally, it should be added that the numerical results presented in this section hold for any phase ϕ in the source signal, $P(0,y)=\sin(y+\phi)$.

B. Gaussian stochastic process with broadband spectrum

Consider now a Gaussian stationary and ergodic stochastic process with broadband PSD to be the source condition. Analytical solutions exist only for the lossless case near the source.[11,2] Time-domain algorithms cannot be used, since the source condition does not contain any phase information. The numerical results obtained by the new method are compared with the analytical solution near the source.

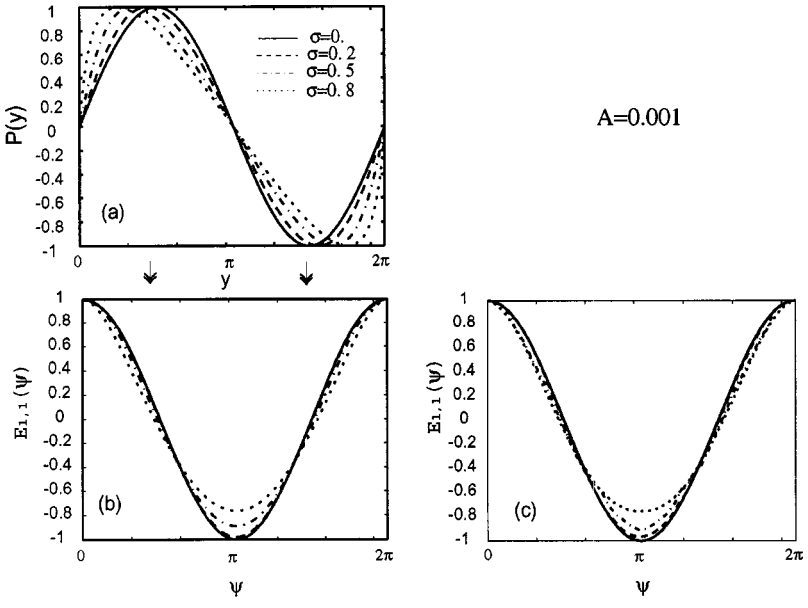


FIG. 7. Comparison of numerical results with known solutions for the case of a sinusoidal source signal $P(y)=\sin y$ and for $A=0.001$. The evolution of the pressure waveform at various propagation distances σ is shown in (a). The corresponding $E_{1,1}(\psi)$ shown in (b) has been obtained by averaging the time-domain results shown in (a). The results shown in (c) have been obtained directly in the joint moments domain by the new method.

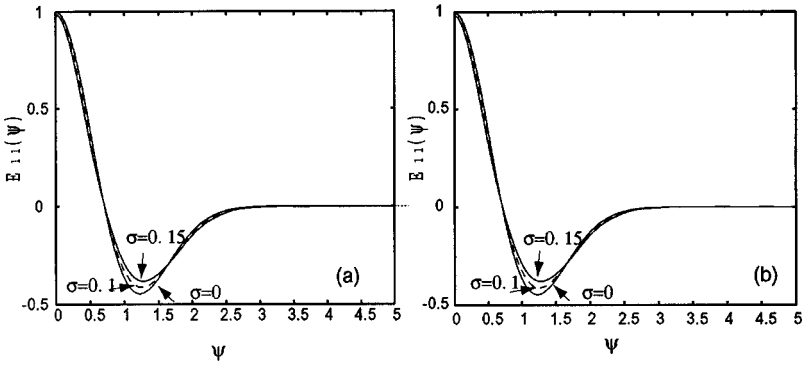


FIG. 8. Comparison of numerical results [shown in (a)] with analytical solutions (Ref. 11) [shown in (b)]. The source condition is a Gaussian, stationary, and ergodic random process with $E_{1,1}(0,\psi)=(1-2\psi^2)e^{-\psi^2}$. The fluid is nondissipative.

Consider a Gaussian stationary and ergodic process with the following power spectral density:

$$S(0,\omega)=\frac{\omega^2}{4\sqrt{\pi}}e^{-\omega^2/4}, \tag{23}$$

where the frequency ω has been normalized by a characteristic frequency ω_0, and the amplitude has been normalized to give unit mean-square value. The inverse Fourier transform of Eq. (23) gives the autocorrelation function at the source

$$E_{1,1}(0,\psi)=(1-2\psi^2)e^{-\psi^2}. \tag{24}$$

An exact analytical solution exists that describes the evolution of the power spectral density in the lossless case. It has been proven[11,2] that near the source, where the waveform distortion is described by the Poisson solution,[3] exact statistical averaging can be performed. The power spectral density is given by the following expression:

$$S(\sigma,\omega)=\frac{e^{-(\omega\sigma)^2}}{2\pi(\omega\sigma)^2}\int_{-\infty}^{+\infty}[e^{(\omega\sigma)^2E_{1,1}(0,\psi)}-1]e^{-i\omega\psi}d\psi. \tag{25}$$

It is emphasized that the analytical solution (25) is valid only near the source. The validity criterion is the size of departure of $E_{1,1}(\sigma,0)$ from its value at the source $E_{1,1}(0,0)=1$. The value of $E_{1,1}(\sigma,0)$ should remain unity, since no losses are included in the model. In the case of Eq. (23) this is true only for $\sigma<0.2$. It should be recalled that σ is the dimensionless distance normalized by $\overline{x}$, which is the lossless plane-wave shock formation distance for a sinusoidal source condition. In Fig. 8 change in the autocorrelation function as a function of σ is shown (a) as given by our numerical computations and (b) as given by the analytical solution. Values of σ are limited to the range where the analytical solution is valid. The propagation step used in the numerical calculations was $\Delta\sigma=0.01$. The agreement between numerical and analytical results is good. Finally, in Fig. 9 the evolution of the corresponding PSD is shown, where the characteristic shift towards high frequencies can be observed.

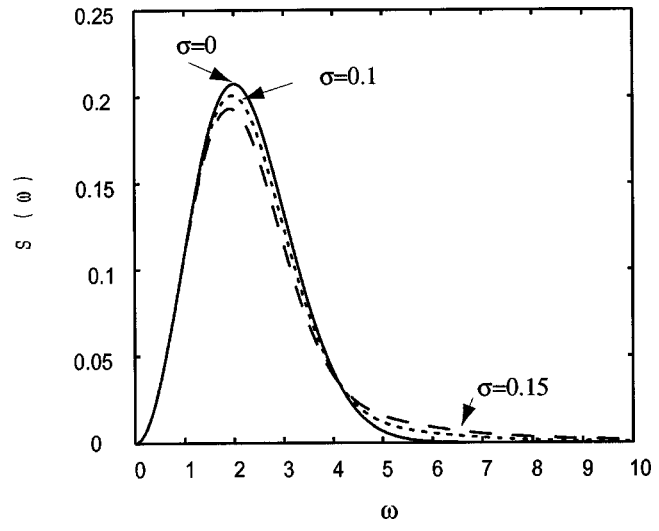


FIG. 9. The evolution of the power spectral density of the Gaussian process shown in Fig. 8.

VII. SUMMARY—FUTURE WORK

A new method to predict the effects of nonlinearity and dissipation on the autocorrelation function of a signal or a stochastic process has been presented. A new equation, termed statistical Burgers equation, has been derived based on the Burgers equation and represents an infinite set of linear equation that describe the evolution of the joint moments of the signal or the process. A method for solving this system of equations by numerical means is presented. It has been shown that the effects of nonlinearity and dissipation can be demonstrated on the autocorrelation function in a similar way as in the time history of the signal. Finally, numerical results from the method have been given and their agreement with time-domain calculations and/or analytical solutions has been demonstrated in the preshock region for two benchmark cases.

In our future work we intend to study the computational limitations that restrict our results in the preshock region and present results in the aftershock region as well. Future work should go beyond benchmark cases and include complex source conditions. Further, we intend to include other linear effects, such as geometrical spreading and attenuation due to relaxation in order to model the problem of intense noise propagation through the real atmosphere.

ACKNOWLEDGMENTS

Professor M. F. Hamilton is gratefully acknowledged for his enlightening ideas and suggestions in the preparation of the present manuscript. The work was supported by the National Aeronautics and Space Administration, the Texas Advanced Technology Program and the F. V. Hunt Postoctoral

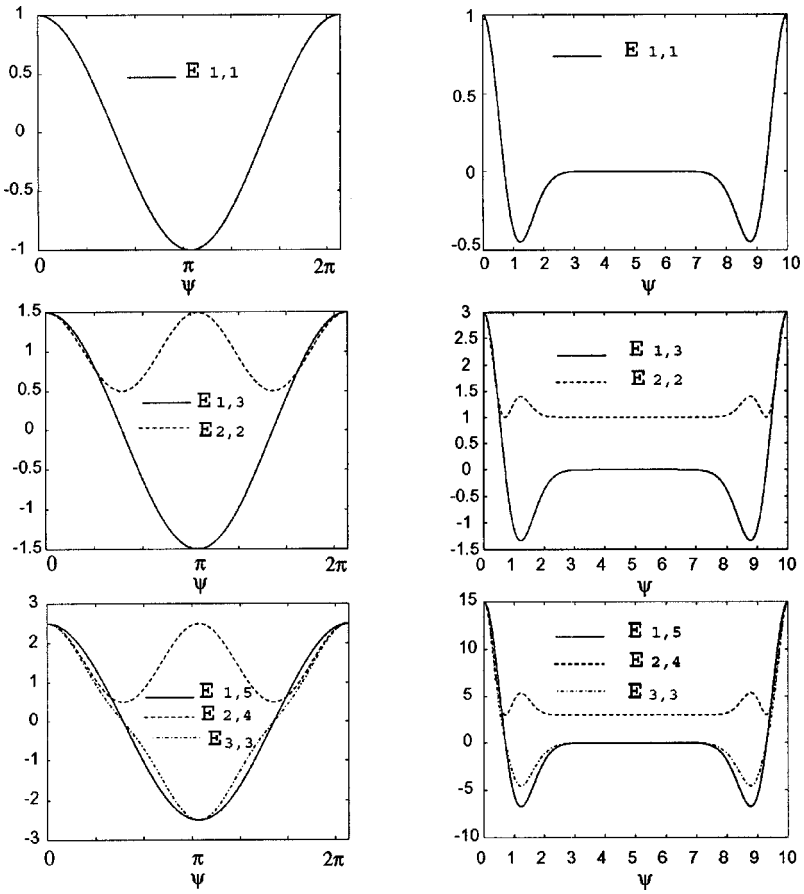


FIG. 10. The joint moments $E_{m,n}$ at the source. Shown on the left are the joint moments [Eq. (A5)], if the source condition is a sinusoidal wave with autocorrelation function $E_{1,1}(0,\psi)=\cos\psi$. Shown on the right are the joint moments [Eq. (A9)], if the source condition is a stochastic Gaussian process with autocorrelation function $E_{1,1}(0,\psi)=(1-2\psi^2)e^{-\psi^2}$. In both cases the joint moments of odd order are identically zero at the source.

Research Fellowship from the Acoustical Society of America. All computations were performed on Sparc 2000 SUN stations provided by the Applied Research Laboratories.

APPENDIX: SOURCE CONDITIONS

In order to apply the method presented here, all the joint moments up to a certain order have to be known at the source. The source condition can be either a single time waveform or a stochastic process. Let us consider the former case. The joint moments at the source can be computed directly by the defining equation [Eq. (9)], if the time history of the source signal is known. If the AF at the source is known, but not the signal itself, then all the joint moments must be expressed in terms of the AF. This can be done (i) if the signal has a Gaussian joint probability density function or (ii) if the signal is sinusoidal (even if the phase of the sinusoidal signal is unknown). Consider now the source condition to be a stochastic process. All the joint moments at the source can be computed from the AF, if the process is a stationary Gaussian process. If the process is not Gaussian, the joint moments have to be given by the statistical description of the process. In general, the method can be applied in all cases the necessary joint moments at the source are given or can be computed.

1. Sinusoidal wave

If the source signal is a sinusoidal wave, then the joint moments can be computed through recursion relations. Compared to computation through the defining equation [Eq. (9)], the computations through recursion relations, if such are available, are preferred, because they are faster and more accurate.

Consider a sinusoidal signal $P(y)=\sqrt{2}\sin(y+\phi)$ that has peak pressure $\sqrt{2}p_0$ normalized by the characteristic pressure p_0 and thus has unit mean-square value. This normalized signal has zero mean

$$E_{1,0}=0, \tag{A1}$$

and normalized autocorrelation function

$$E_{1,1}=\cos\psi=R(\psi). \tag{A2}$$

In that case the following are true:

$$E_{m,0}=E_{m,n}(0)=\begin{cases} 0 & \text{if } m \text{ is odd} \\ \dfrac{1\cdot 3\cdots(m-1)}{\left(\dfrac{m}{2}\right)!} & \text{if } m \text{ is even.} \end{cases} \tag{A3}$$

Furthermore, since the signal is stationary the following relations hold: $E_{m,0}=E_{0,m}=E_{m,n}(0)$. The result in Eq. (A3) can be extended for the function $E_{m,n}(\psi)$ if both m and n are nonzero

$$E_{m,n}(\psi)=\begin{cases} 0 & \text{if } m+n \text{ is odd} \\ \neq 0 & \text{if } m+n \text{ is even.} \end{cases} \tag{A4}$$

It can be shown that for $m\geqslant 2$ and $n\geqslant 1$ the following recursion relation holds:

$$\frac{m+n}{2}E_{m,n}=nR(\psi)E_{m-1,n-1}+(m-1)E_{m-2,n}. \quad \text{(A5)}$$

Equation (A5), together with Eq. (A1), can create all joint moments $E_{m,n}$. The symmetry of $E_{m,n}$ [Eq. (10)] can be used to reduce the number of calculations. It should be noted that the amplitude of the joint moments increases very fast as the order of the moment increases.

The recursion equation (A5) shows that $E_{m,n}(\psi)$ is an even function, if $m+n$ is an even number, and $E_{m,n}(\psi)=0$, if $m+n$ is an odd number. Therefore, at the source

$$E_{m,n}(\psi)=\begin{cases}0 & \text{if } m+n \text{ is odd}\\ E_{m,n}(-\psi)=E_{n,m}(\psi) & \text{if } m+n \text{ is even.}\end{cases} \quad \text{(A6)}$$

Figure 10 shows the shape of the joint moments at the source.

As the signal travels along the propagation path, it undergoes nonlinear distortion. Equation (15) shows that the initially zero-valued moments of odd order become nonzero. More specifically, Eq. (15) reveals that the joint moments of even order remain even functions throughout the propagation, and that the joint moments of odd order (zero at the source) become and remain odd functions

$$E_{m,n}(\sigma,-\psi)=\begin{cases}-E_{m,n}(\sigma,\psi) & \text{if } m+n \text{ is odd}\\ E_{m,n}(\sigma,\psi) & \text{if } m+n \text{ is even.}\end{cases} \quad \text{(A7)}$$

It can be proven that Eq. (A7) holds for any signal with odd symmetry, which maintains its odd symmetry throughout the propagation.

It should finally be noted that all the results presented in this section for a single timeform $P(y)=\sqrt{2}\sin(y+\phi)$ hold also for the stationary stochastic process $\mathbf{P}(y)=\sqrt{2}\sin(y+\mathbf{\Phi})$, where $\mathbf{\Phi}$ is a random variable uniformly distributed in $(-\pi, +\pi)$.

2. Gaussian stochastic process

Consider a stationary and ergodic random process $\mathbf{P}(y)$ with zero mean and normalized to unit mean-square value. As in the case of a sinusoidal signal the latter specifies the characteristic pressure p_0. The normalized process has zero mean

$$E_{1,0}=0,$$

and normalized autocorrelation function

$$E_{1,1}=R(\psi).$$

If the process is Gaussian, the following is true:

$$E_{m,0}=\begin{cases}0 & \text{if } m \text{ is odd}\\ 1\cdot 3\cdots(m-1) & \text{if } m \text{ is even.}\end{cases} \quad \text{(A8)}$$

Because the signal is stationary, the following relations hold: $E_{m,0}=E_{0,m}=E_{m,n}(0)$. Furthermore, because the process is Gaussian all joint moments $E_{m,n}$ can be determined in terms of the autocorrelation function $E_{1,1}$. More specifically, it can be shown[29] that for $m\geqslant 2$ and $n\geqslant 1$ the following recursion relation holds:

$$E_{m,n}=nR(\psi)E_{m-1,n-1}+(m-1)E_{m-2,n}. \quad \text{(A9)}$$

Equation (A9) is valid only at the source. It cannot be used for the later stages of propagation, because the process undergoes nonlinear propagation distortion and becomes non-Gaussian. However, the process remains stationary and ergodic throughout the propagation.[30,31]

Equation (A9) is similar in form with Eq. (A5) for the sinusoidal case. As in the case of a sinusoidal wave, Eqs. (A6) and (A7) specify the symmetry conditions of the joint moments at the source and throughout the propagation, respectively. In the right column of Fig. 10 the joint moments at source are presented for a Gaussian stochastic process with autocorrelation function

$$R(\psi)=(1-2\psi^2)e^{-\psi^2}.$$

It can be observed that the amplitude of the joint moments increases with the order of the moment faster than it does in the case of a sinusoidal wave (shown in the same figure on the left).

It should be noted that stationary noise signals with Gaussian joint probability density function can be treated in the same way.

[1] M. J. Lighthill, Waves in Fluids (Cambridge University Press, Cambridge, England, 1980).
[2] O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Plenum, New York, 1977), Chap. 10.
[3] *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, San Diego, 1998).
[4] J. D. Cole, "On a quasi-linear parabolic equation occurring in aerodynamics," Q. Appl. Math. **9**, 225–236 (1951).
[5] E. Hopf, "The partial differential equation $u_t+uu_x=\mu_{xx}$," Commun. Pure Appl. Math. **3**, 201–230 (1950).
[6] F. H. Fenlon, "A recursive procedure for computing the nonlinear spectral interactions of progressive finite-amplitude waves in nondispersive fluids," J. Acoust. Soc. Am. **50**, 1299–1312 (1971).
[7] F. M. Pestorius and D. T. Blackstock, "Propagation of finite-amplitude noise," in *Finite-Amplitude Wave Effects in Fluids*, edited by L. Bjorno (IPC Science and Technology, London, England, 1974).
[8] Y.-S. Lee and M. F. Hamilton, "Time-domain modeling of pulsed finite-amplitude sound beams," J. Acoust. Soc. Am. **97**, 906–916 (1995).
[9] R. O. Cleveland, M. F. Hamilton, and D. T. Blackstock, "Time modeling of finite-amplitude sound in relaxing fluids," J. Acoust. Soc. Am. **99**, 3312–3318 (1996).
[10] F. M. Pestorius, S. W. Williams, and D. T. Blackstock, "Effect of nonlinearity on noise propagation," in 2nd Interagency Symposium on University Research in Transportation Noise, North Carolina State University, Raleigh, 1974.
[11] O. V. Rudenko and A. S. Chirkin, "Nonlinear transformation of random wavefield spectra," Dokl. Akad. Nauk SSSR **214**, 1045–1048 (1974).
[12] O. V. Rudenko, "Interactions of intense noise waves," Sov. Phys. Usp. **29**, 620–641 (1986).
[13] D. G. Crighton and S. Bashforth, "Nonlinear propagation of broadband jet noise," Institute of Aeronautics and Astronautics, 6th Aeroacoustic Conference, Paper 80-1039 (1980).
[14] S. Gurbatov, A. Malakhov, and A. Saichev, *Nonlinear Random Waves and Turbulence in Nondispersive Media: Waves, Rays and Particles* (Manchester University Press, Manchester-New York, 1991), Chaps. 1 and 4.
[15] J. M. Burgers, "A mathematical model illustrating the theory of turbulence," in *Advances in Applied Mechanics*, edited by R. von Mises and T. von Karman (Academic, New York, 1948), Vol. 1, pp. 171–199.
[16] R. H. Kraichnan, "The structure of isotropic turbulence at very high Reynolds numbers," J. Fluid Mech. **5**, 497–543 (1959).
[17] L. S. G. Kovasznay, "Spectrum of locally isotropic turbulence," J. Aeronaut. Sci. **15**, 745–753 (1948).

[18] W. Heisenberg, ''On the theory of statistical and isotropic turbulence,'' Proc. R. Soc. London, Ser. A **195**, 402–406 (1948).
[19] Y. H. Pao, ''Structure of turbulent velocity and scalar fields at large wave numbers,'' Phys. Fluids **8**, 1063–1075 (1965).
[20] J. J. Walton, ''Integration of the Langrangian-History approximation to the Burgers' equation,'' Phys. Fluids **13**, 1634–1635 (1970).
[21] M. Mond and G. Knorr, ''Higher corrections to the direct interaction approximation in turbulence theory,'' Phys. Fluids **23**, 1306–1310 (1980).
[22] C. L. Morfey, ''Nonlinear propagation of jet noise in the atmosphere,'' Royal Aircraft Establishment, Technical Report 80004 (1980).
[23] G. P. Howell and C. L. Morfey, ''Nonlinear propagation of broadband noise signals,'' J. Sound Vib. **114**, 189–201 (1987).
[24] D. T. Jeng, R. Foester, S. Haaland, and W. C. Meecham, ''Statistical initial-value problem for Burgers' model equation of turbulence,'' Phys. Fluids **9**, 2114–2121 (1966).
[25] Y. J. Ogura, ''Statistical hydrodynamics for burgers turbulence,'' J. Fluid Mech. **16**, 38–41 (1963).
[26] A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1991), Chaps. 7 and 10.
[27] D. A. Anderson, J. C. Tannehill, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer* (McGraw-Hill, New York, 1984).
[28] D. T. Blackstock, ''Connection between the Fay and Fubini solutions for plane sound waves of finite amplitude,'' J. Acoust. Soc. Am. **39**, 1019–1026 (1966).
[29] Y. K. Lin, *Probabilistic Theory of Structural Dynamics* (McGraw-Hill, New York, 1967).
[30] O. A. Vasil'eva, A. A. Karabutov, E. A. Lapshin, and O. V. Rudenko, *Interaction of One-Dimensional Waves in Nondispersive Media* (Moscow University Press, Moscow, 1983).
[31] O. A. Vasil'eva, *Numerical Analysis using* FORTRAN, edited by V. V. Voevodin and V. A. Morozov (Moscow University Press, Moscow, 1983), p. 47.

Nonlinear dynamics of a gas bubble in an incompressible elastic medium

Stanislav Y. Emelianov
Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712-1084

Mark F. Hamilton,[a] Yurii A. Ilinskii, and Evgenia A. Zabolotskaya
Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712-1063 and Applied Research Laboratories, The University of Texas at Austin, Austin, Texas 78713-8029

(Received 23 May 2003; accepted for publication 4 September 2003)

A nonlinear model in the form of the Rayleigh–Plesset equation is developed for a gas bubble in an essentially incompressible elastic medium such as a tissue or rubberlike medium. Two constitutive laws for the elastic medium are considered: the Mooney potential, and Landau's expansion of the strain energy density. These two constitutive laws are compared at quadratic order to obtain a relation between their respective elastic constants. Attention is devoted to the relative importance of shear stress on the bubble dynamics, allowing for the equilibrium gas pressure in the bubble to differ substantially from the pressure at infinity. The model for the bubble motion is approximated to quadratic order to assess the importance of shear stress in the surrounding medium relative to that of the gas pressure in the bubble. Relations are derived for the value of the shear wave speed at which the two contributions are comparable, which provide an assessment of when shear stress in the surrounding medium must be taken into account when modeling bubble dynamics. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1621858]

PACS numbers: 43.25.Yw, 43.25.Dc [AJS] Pages: 581–588

I. INTRODUCTION

Biomedical applications utilizing microbubbles are the principal motivation for recent theoretical studies of bubble dynamics in viscoelastic media (see Refs. 1 and 2 for reviews of this literature). These investigations are due to the recognition that non-Newtonian fluids often surround the bubble. Bubbles encapsulated in thin elastic shells, manufactured for use as contrast agents and being considered for drug or gene delivery vehicles, are also modeled using constitutive relations for viscoelastic media.[3]

In the present investigation the medium surrounding a spherical, gas-filled bubble is described using standard methods employed in the theory of elasticity. The main assumptions are that the medium is isotropic and incompressible, but with nonlinearity taken into account. The oscillation of a spherical cavity in an incompressible medium is a classical problem in the theory of elasticity. The derivation of the natural frequency in the linear approximation is reviewed by Landau and Lifshitz,[4] early measurements of this frequency are reported by Meyer *et al.*,[5] and a dynamical equation for arbitrarily large radial motion is derived in Rayleigh–Plesset form by Eringen and Suhubi.[6] These analyses have been extended to include gas in the cavity as well as other factors associated with bubble oscillations by Alekseev and Rybak[7] in the linear approximation, and by Ostrovsky[8,9] for oscillations of finite amplitude.

The analysis presented below, although similar to that of Ostrovsky,[8,9] differs in several respects. A fundamental distinction involves the choice of equilibrium radius about which bubble wall displacement is analyzed. Here we consider the equilibrium radius in the nonuniformly prestressed state that occurs when the equilibrium gas pressure in the bubble differs from the pressure at infinity. Ostrovsky takes as his reference the unstressed equilibrium state that occurs only when these two pressures are equal. The present approach is necessary for describing oscillations of pressurized cavities or of bubbles created by vaporization. Another distinction is that the present analysis is performed in terms of bubble radius, rather than bubble volume, to facilitate comparison with the Rayleigh–Plesset equation. Because both the differential equation describing the bubble oscillation and the relation between radius and volume are nonlinear, the corresponding analytical results expressed in terms of radius and volume can differ considerably in appearance. Finally, explicit results developed by Ostrovsky are based on the Mooney[10] potential for the strain energy density in an incompressible medium. Media described by this constitutive relation are often referred to as Mooney[10]–Rivlin[11] materials. This relation is used frequently for rubberlike media, and it is finding increasing application as a model for tissue. Alternatively, Landau's expansion[4] of the strain energy density is more general, and the associated elastic moduli are more widely understood and reported in the literature. Both constitutive relations are considered below, and a relation connecting their respective elastic moduli is derived.

In Sec. II, a general equation for the radial motion of a spherical, gas-filled bubble in an incompressible elastic medium is developed using Lagrangian mechanics. An integral expression for the contribution due to shear stress in the elastic medium is derived in Sec. III for strain energy density functions expressed in terms of Green's deformation tensor. The integral is evaluated and investigated in Sec. IV for elastic media described by Mooney's constitutive relation. In

[a]Electronic mail: hamilton@mail.utexas.edu

Sec. V, the contribution due to shear stress is investigated using Landau's expansion of the strain energy density in terms of the Lagrangian strain tensor. In Sec. VI, the equation of motion is approximated at quadratic order in terms of radial displacement relative to an equilibrium position corresponding to an equilibrium gas pressure that may differ substantially from the pressure at infinity. This result permits an assessment of the relative importance of potential energy stored in compression of the gas and in shear deformation of the elastic medium. As an example it is shown that for soft tissue that is unstressed in static equilibrium, shear stress must be taken into account at both linear and quadratic order when the shear wave speed in the surrounding medium is of order 10 m/s or more.

The purpose of the present article is to provide a formalism for describing nonlinear bubble oscillations in media having small but finite shear moduli, and to assess the relative influence of shear stress on the resulting bubble dynamics. Numerical solutions for bubble radius as a function of time are not presented. We also do not consider additional factors that are commonly taken into account in the Rayleigh–Plesset equation, such as viscosity, radiation loss, and acoustical excitation. Such numerical solutions and the inclusion of other effects shall be reported elsewhere.

II. EQUATION OF MOTION

Here we develop the dynamical equation for radial oscillation of a spherical bubble surrounded by an incompressible elastic medium. The assumption of incompressibility supposes the condition that the bubble radius $R(t)$ is very small in comparison with the wavelength of a compressional disturbance, or $kR_0 \ll 1$, where R_0 here is a characteristic bubble radius, and $k=\omega/c_l$ is the compressional wave number in the surrounding medium associated with motion of the bubble at angular frequency ω and for longitudinal wave speed c_l. If the characteristic frequency is taken to be the resonance of an empty spherical cavity in an incompressible medium,[4] $\omega=2c_t/R_0$, where c_t is the transverse (shear) wave speed, then the inequality corresponds to $c_t \ll c_l$. With $c_t=(\mu/\rho)^{1/2}$ and $c_l=[(K+\frac{4}{3}\mu)/\rho]^{1/2}$, where μ is the shear modulus, K the bulk modulus, and ρ the density of the elastic medium, the assumption of incompressibility in the present context corresponds to $\mu \ll K$.

The bubble is assumed to contain gas having pressure P_g in static equilibrium. The pressure at the outer boundary of the elastic medium, assumed infinitely far away from the bubble, is designated P_∞. It is important to make a distinction between equilibrium bubble radii for $P_g=P_\infty$ and $P_g \neq P_\infty$. The initial undeformed state is taken to be the condition of equilibrium for which $P_g=P_\infty$. In this state the shear stress is uniform throughout the elastic medium, and the equilibrium bubble radius is designated R_0 [see Fig. 1(a)]. For $P_g \neq P_\infty$ the shear stress is nonuniform, the deviation in stress from its state at infinity is most pronounced in the neighborhood of the bubble, and in this case the equilibrium bubble radius is designated R_1 [see Fig. 1(b)].

Lagrange's equation for the radial pulsation of the bubble is

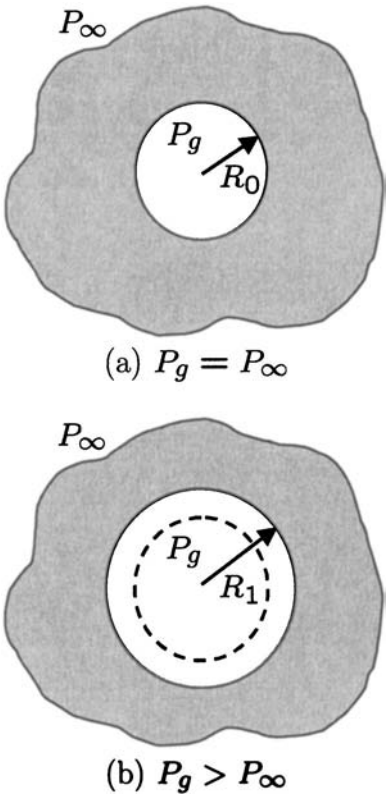


FIG. 1. Depiction of equilibrium bubble radii for (a) $P_g=P_\infty$ and (b) $P_g>P_\infty$.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{R}}\right)-\frac{\partial L}{\partial R}=0, \tag{1}$$

where the bubble radius R is taken as the generalized coordinate. The Lagrangian L is

$$L=T-U, \tag{2}$$

where T is the kinetic energy and U the potential energy of the entire system. Assuming the kinetic energy of the gas in the bubble to be negligible in comparison with that in the elastic medium, we have

$$T=\frac{\rho}{2}\int_{r=R}^{\infty} v^2\, dV, \tag{3}$$

where ρ is the (constant) density and v is the radial particle velocity, respectively, in the medium, and $dV=4\pi r^2\, dr$ is the volume element in terms of the Eulerian radial coordinate r. From continuity we have, for an incompressible medium,

$$v=\dot{R}\frac{R^2}{r^2}, \tag{4}$$

the substitution of which in Eq. (3) yields

$$T=2\pi\rho\dot{R}^2R^3 \tag{5}$$

for the kinetic energy in the system.

The potential energy is the sum of the internal energy of the gas in the bubble (U_g), the energy of deformation in the elastic medium (U_e), and the contribution due to work done by pressure on the outer surface of the elastic medium at infinity (U_∞):

$$U = U_g + U_\infty + U_e. \tag{6}$$

For an ideal gas $U_g = PV/(\gamma - 1)$, where P is pressure, $V = \frac{4}{3}\pi R^3$ is bubble volume, and γ is the ratio of specific heats. For an adiabatic gas $P/P_g = (V_1/V)^\gamma$, where $V_1 = \frac{4}{3}\pi R_1^3$ is the equilibrium bubble volume, and thus

$$U_g = \frac{P_g V_1}{\gamma - 1}\left(\frac{R_1}{R}\right)^{3(\gamma-1)}. \tag{7}$$

The work done by the pressure P_∞ on the outer surface of the elastic medium gives $U_\infty = P_\infty(V - V_1)$, and therefore

$$U_\infty = P_\infty V_1\left[\left(\frac{R}{R_1}\right)^3 - 1\right]. \tag{8}$$

For an adiabatic system U_e is the internal energy of the elastic medium. The calculation of U_e requires integration of an elastic energy density function $\mathcal{E}_e$ expressed in terms of the Lagrangian coordinate $\tilde{r}$ (distinguished from Eulerian coordinates with over tildes):

$$U_e = \int_{\tilde{r}=R_0}^{\infty} \mathcal{E}_e \, d\tilde{V}. \tag{9}$$

Since $\mathcal{E}_e$ is a function of invariants that can be expressed through $\tilde{r}$ and R, as described below, one may write $\mathcal{E}_e = \mathcal{E}_e(\tilde{r}, R)$, and thus

$$U_e = 4\pi \int_{R_0}^{\infty} \mathcal{E}_e \tilde{r}^2 \, d\tilde{r}. \tag{10}$$

The Lagrangian L is now determined apart from an explicit expression for U_e, and the evaluation of Eq. (1) yields

$$R\ddot{R} + \frac{3}{2}\dot{R}^2 = \frac{1}{\rho}\left[P_g\left(\frac{R_1}{R}\right)^{3\gamma} - P_\infty - P_e(R)\right], \tag{11}$$

where

$$P_e(R) = \frac{1}{4\pi R^2}\frac{\partial U_e}{\partial R} \tag{12}$$

is an effective pressure resulting from shear stress. The defining relation for the equilibrium radius R_1 is obtained by setting $R = R_1$ and $\dot{R} = \ddot{R} = 0$ in Eq. (11):

$$P_e(R_1) = P_g - P_\infty. \tag{13}$$

When the bubble is surrounded by a liquid rather than an elastic medium the shear stress vanishes, such that $U_e = 0$, $P_g = P_\infty$, and a form of the Rayleigh–Plesset equation is recovered.[12] The problem has thus been reduced to elastic energy calculations.

III. ELASTIC ENERGY

The elastic energy density is frequently expressed in terms of invariants of Green's deformation tensor,[6]

$$G_{i'k'} = \frac{\partial x_l}{\partial \tilde{x}_{i'}}\frac{\partial x_l}{\partial \tilde{x}_{k'}}. \tag{14}$$

In addition to overtildes designating Lagrangian coordinates, primed subscripts designate their indices. Unprimed subscripts are indices for the Eulerian coordinates. The principal invariants of $G_{i'k'}$ are

$$I_1 = G_{i'i'} = \nu_1^2 + \nu_2^2 + \nu_3^2, \tag{15}$$

$$I_2 = \tfrac{1}{2}(G_{i'i'}^2 - G_{i'k'}G_{k'i'}) = \nu_1^2\nu_2^2 + \nu_2^2\nu_3^2 + \nu_1^2\nu_3^2, \tag{16}$$

$$I_3 = \det G_{i'k'} = \nu_1^2\nu_2^2\nu_3^2, \tag{17}$$

where ν_1^2, ν_2^2, and ν_3^2 are the principal values. Since the determinant of Eq. (14) equals the square of the Jacobian for Eulerian and Lagrangian coordinate systems, and since this Jacobian is unity when volume elements do not change, one obtains

$$I_3 = 1 \tag{18}$$

for incompressible media. On the basis of this relation, corresponding to $\nu_1^2\nu_2^2\nu_3^2 = 1$, it is convenient to rewrite Eq. (16) as

$$I_2 = G_{i'i'}^{-1} = \frac{1}{\nu_1^2} + \frac{1}{\nu_2^2} + \frac{1}{\nu_3^2}. \tag{19}$$

We thus proceed to express the elastic energy of an incompressible medium in terms of the principal invariants I_1 and I_2 defined by Eqs. (15) and (19), respectively.

The principal values are now calculated in spherical coordinates, for which

$$\frac{x_i}{\tilde{x}_i} = \frac{r}{\tilde{r}}, \quad \frac{\partial \tilde{r}}{\partial \tilde{x}_i} = \frac{\tilde{x}_i}{\tilde{r}}, \tag{20}$$

and thus

$$\frac{\partial x_i}{\partial \tilde{x}_k} = \frac{\tilde{x}_i\tilde{x}_k}{\tilde{r}^2}\left(\frac{\partial r}{\partial \tilde{r}} - \frac{r}{\tilde{r}}\right) + \frac{r}{\tilde{r}}\delta_{ik}. \tag{21}$$

Calculation of the principal values is simplified by evaluating $G_{i'k'}$ at the point $(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3) = (0, 0, \tilde{r})$, for which Eq. (21) yields

$$\frac{\partial x_i}{\partial \tilde{x}_k} = \begin{pmatrix} r/\tilde{r} & 0 & 0 \\ 0 & r/\tilde{r} & 0 \\ 0 & 0 & \partial r/\partial \tilde{r} \end{pmatrix}, \tag{22}$$

and $G_{i'k'}$ assumes the following diagonalized form:

$$G_{i'k'} = \begin{pmatrix} (r/\tilde{r})^2 & 0 & 0 \\ 0 & (r/\tilde{r})^2 & 0 \\ 0 & 0 & (\partial r/\partial \tilde{r})^2 \end{pmatrix}. \tag{23}$$

The principal values are observed to be

$$\nu_1^2 = \nu_2^2 = \left(\frac{r}{\tilde{r}}\right)^2, \quad \nu_3^2 = \left(\frac{\partial r}{\partial \tilde{r}}\right)^2 = \left(\frac{\tilde{r}}{r}\right)^4, \tag{24}$$

where the last expression for ν_3^2 follows from Eqs. (17) and (18):

$$\frac{\partial r}{\partial \tilde{r}} = \frac{\tilde{r}^2}{r^2}. \tag{25}$$

Equations (15) and (19) become[6]

$$I_1 = 2\frac{r^2}{\tilde{r}^2} + \frac{\tilde{r}^4}{r^4}, \quad I_2 = 2\frac{\tilde{r}^2}{r^2} + \frac{r^4}{\tilde{r}^4}, \tag{26}$$

and for an incompressible medium, from integration of Eq. (25),

$$r^3 = \tilde{r}^3 + R^3 - R_0^3. \tag{27}$$

We may thus write $I_{1,2} = I_{1,2}(\tilde{r}, R)$. From Eq. (10)

$$U_e = 4\pi \int_{R_0}^{\infty} \mathcal{E}_e(I_1, I_2)\tilde{r}^2\, d\tilde{r}, \tag{28}$$

and therefore given the function $\mathcal{E}_e(I_1, I_2)$ one can calculate U_e as an explicit function of R.

Equation (12) requires the derivative $\partial U_e/\partial R$, which can be expressed as

$$\frac{\partial U_e}{\partial R} = 4\pi \int_{R_0}^{\infty} \left(\frac{\partial \mathcal{E}_e}{\partial I_1}\frac{\partial I_1}{\partial R} + \frac{\partial \mathcal{E}_e}{\partial I_2}\frac{\partial I_2}{\partial R} \right) \tilde{r}^2\, d\tilde{r} \tag{29}$$

on the basis of Eq. (28). From Eqs. (26) and (27)

$$\frac{\partial I_1}{\partial R} = \frac{2}{\tilde{r}^2}\frac{\partial (r^2)}{\partial R} + \tilde{r}^4 \frac{\partial (r^{-4})}{\partial R} = 4\frac{R^2}{\tilde{r}^3}\left(\frac{\tilde{r}}{r} - \frac{\tilde{r}^7}{r^7} \right), \tag{30}$$

$$\frac{\partial I_2}{\partial R} = 2\tilde{r}^2 \frac{\partial (r^{-2})}{\partial R} + \frac{1}{\tilde{r}^4}\frac{\partial (r^4)}{\partial R} = 4\frac{R^2}{\tilde{r}^3}\left(\frac{r}{\tilde{r}} - \frac{\tilde{r}^5}{r^5} \right), \tag{31}$$

substitution of which in Eq. (29) yields

$$\frac{\partial U_e}{\partial R} = 16\pi R^2 \int_{R_0}^{\infty} \left(\frac{\partial \mathcal{E}_e}{\partial I_1} + \frac{r^2}{\tilde{r}^2}\frac{\partial \mathcal{E}_e}{\partial I_2} \right) \left(\frac{\tilde{r}}{r} - \frac{\tilde{r}^7}{r^7} \right) \frac{d\tilde{r}}{\tilde{r}}. \tag{32}$$

Equation (12) thus becomes

$$P_e(R) = 4\int_{R_0}^{\infty} \left(\frac{\partial \mathcal{E}_e}{\partial I_1} + \frac{r^2}{\tilde{r}^2}\frac{\partial \mathcal{E}_e}{\partial I_2} \right) \left(\frac{\tilde{r}}{r} - \frac{\tilde{r}^7}{r^7} \right) \frac{d\tilde{r}}{\tilde{r}}, \tag{33}$$

which is a general relation for the elastic force per unit area associated with radial displacement in an incompressible elastic medium described by an elastic energy density expressed in terms of invariants of Green's deformation tensor.

Together, Eqs. (11) and (33) are formally equivalent to Eq. (3.4.6) of Eringen and Suhubi[6] for a spherical shell when the outer radius is taken to be infinite. The results differ in appearance only because the integration variable in the latter equation is the Eulerian coordinate r, rather than the Lagrangian coordinate $\tilde{r}$ as in Eq. (33). Equations (11) and (33) also correspond to Eq. (6) on p. 348 of Lurie[13] for a spherical shell when its outer radius is taken to be infinite. The latter is the starting point for Ostrovsky.[8] Lurie's dynamical equation differs from the present formulation because it is expressed in terms of the volume deviation $R^3 - R_0^3$ rather than the bubble radius R, and the elastic medium is assumed to be in a vaccuum. Ostrovsky[8,9] subsequently accounted for finite pressures in the interior and exterior of the elastic medium. However, the motivation for the present work is not to emphasize differences with earlier formulations, but rather to interpret $P_e(R)$ in the context of gas bubbles described by the Rayleigh–Plesset equation.

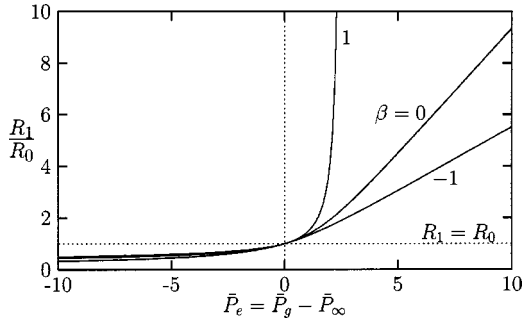


FIG. 2. Relative equilibrium bubble radius R_1/R_0 as a function of $P_e = P_g - P_\infty$ for several values of the parameter β in Mooney's constitutive relation. The bars over pressure variables on the horizontal axis indicate dimensionless quantities formed by dividing by shear modulus, e.g., $\bar{P}_g = P_g/\mu$.

IV. MOONEY'S CONSTITUTIVE RELATION

By far, the most widely used model equation for the elastic energy density in an incompressible medium is Mooney's constitutive relation,[10,13]

$$\mathcal{E}_e = \frac{\mu}{4}[(1+\beta)(I_1 - 3) + (1-\beta)(I_2 - 3)], \tag{34}$$

where μ is the shear modulus, β is a fitting parameter referred to by Mooney as the coefficient of asymmetry,[10] which must lie in the range $-1 \leqslant \beta \leqslant 1$ for $\mathcal{E}_e$ to be positive,[13] and $I_{1,2}$ are the first and second principal invariants of Green's deformation tensor. The derivatives with respect to the invariants are

$$\frac{\partial \mathcal{E}_e}{\partial I_1} = \frac{\mu}{4}(1+\beta), \quad \frac{\partial \mathcal{E}_e}{\partial I_2} = \frac{\mu}{4}(1-\beta), \tag{35}$$

and Eq. (33) becomes

$$P_e(R) = \mu \int_{R_0}^{\infty} \left[(1+\beta)\left(\frac{\tilde{r}}{r} - \frac{\tilde{r}^7}{r^7} \right) + (1-\beta)\left(\frac{r}{\tilde{r}} - \frac{\tilde{r}^5}{r^5} \right) \right] \frac{d\tilde{r}}{\tilde{r}}. \tag{36}$$

To evaluate the integral it is convenient to introduce the dimensionless coordinates

$$x = \frac{R_0}{\tilde{r}}, \quad y = \frac{R_0}{r}, \quad \zeta = \frac{R}{R_0}, \tag{37}$$

which are related according to Eq. (27) as

$$y = (x^{-3} + \zeta^3 - 1)^{-1/3}. \tag{38}$$

Equation (36) thus becomes

$$P_e(\zeta) = \mu \int_0^1 \left[(1+\beta)\left(\frac{y}{x^2} - \frac{y^7}{x^8} \right) + (1-\beta)\left(\frac{1}{y} - \frac{y^5}{x^6} \right) \right] dx. \tag{39}$$

With an explicit expression for P_e, Eq. (13) can be solved to determine the equilibrium radius R_1 of a gas bubble as a function of the equilibrium gas pressure P_g in relation to the external pressure at infinity, P_∞. The results of these calculations, for $\beta = 1$, 0, and -1, are displayed in Fig. 2. The bars over the pressure variables on the horizontal axis indicate dimensionless quantities formed by dividing by μ.

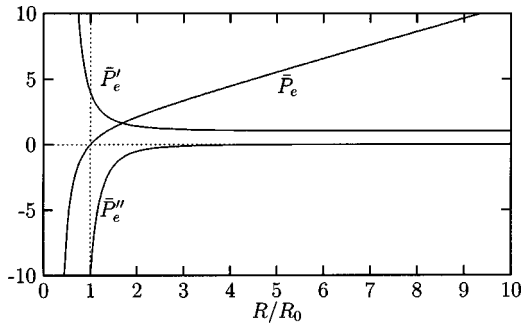


FIG. 3. Dependence of $\bar{P}_e = P_e/\mu$ and its first two derivatives on R/R_0 as defined by Eqs. (39), (43), and (44).

The result $R_1 = R_0$ is obtained for $P_g = P_\infty$ regardless of the value of β, in agreement with the definition of R_0. It is also seen that, independent of β, one obtains $R_1 < R_0$ for $P_g < P_\infty$, and $R_1 > R_0$ for $P_g > P_\infty$, as expected. For $R_1 \simeq R_0$, the parameter β can be related to Landau's elastic constants[4] through Eq. (64) below.

A useful analytical approximation of Eq. (39) is now obtained as an expansion about the equilibrium radius $R = R_1$. Defining

$$\zeta_1 = R_1/R_0, \tag{40}$$

one may write

$$P_e(\zeta) = P_e(\zeta_1) + (\zeta - \zeta_1)P_e'(\zeta_1) + \tfrac{1}{2}(\zeta - \zeta_1)^2 P_e''(\zeta_1) \tag{41}$$

to quadratic order in $(\zeta - \zeta_1)$, where the primes indicate derivatives with respect to ζ. From Eq. (38)

$$\frac{\partial}{\partial \zeta} = -\zeta^2 y^4 \frac{\partial}{\partial y}, \tag{42}$$

and thus, from Eq. (39),

$$P_e'(\zeta) = \mu\zeta^2 \int_0^1 \left[(1+\beta)\left(7\frac{y^{10}}{x^8} - \frac{y^4}{x^2} \right) + (1-\beta)\left(y^2 + 5\frac{y^8}{x^6} \right) \right] dx, \tag{43}$$

$$P_e''(\zeta) = \mu\zeta^4 \int_0^1 \left[(1+\beta)\left(4\frac{y^7}{x^2} - 70\frac{y^{13}}{x^8} \right) - (1-\beta)\left(2y^5 + 40\frac{y^{11}}{x^6} \right) \right] dx + \frac{2}{\zeta} P_e'(\zeta). \tag{44}$$

Normalized values of P_e and its first two derivatives, based on the above integrals, are displayed in Fig. 3 for $\beta = 0$.

The integrals in Eqs. (43) and (44) can be evaluated analytically for $\zeta = 1$ (i.e., for $R = R_0$), because $y = x$ from Eq. (38), and thus

$$P_e'(1) = 4\mu, \quad P_e''(1) = -2\mu(5+2\beta). \tag{45}$$

Taking $\zeta_1 = 1$ in Eq. (41) corresponds to the equilibrium state for which $R_1 = R_0$. This case occurs when $P_g = P_\infty$. With $\zeta - \zeta_1 = \delta R/R_0$, where

$$\delta R = R - R_0, \tag{46}$$

Eq. (41) becomes, with $P_e(R_0) = 0$ from Eq. (11),

$$P_e = 4\mu \frac{\delta R}{R_0} - (5+2\beta)\mu \left(\frac{\delta R}{R_0} \right)^2, \tag{47}$$

which permits the determination of β in terms of the elastic constants in Landau's constitutive relation, considered next.

V. LANDAU'S CONSTITUTIVE RELATION

Landau[4] expresses the elastic energy density in terms of invariants of the Lagrangian strain tensor

$$u_{i'k'} = \frac{1}{2}\left(\frac{\partial u_i}{\partial \tilde{x}_{k'}} + \frac{\partial u_k}{\partial \tilde{x}_{i'}} + \frac{\partial u_l}{\partial \tilde{x}_{i'}} \frac{\partial u_l}{\partial \tilde{x}_{k'}} \right), \tag{48}$$

where u_i is the particle displacement vector. Instead of the principal invariants of this tensor he uses the alternative set of invariant

$$L_1 = u_{i'i'} = \eta_1 + \eta_2 + \eta_3, \tag{49}$$

$$L_2 = u_{i'k'}u_{k'i'} = \eta_1^2 + \eta_2^2 + \eta_3^2, \tag{50}$$

$$L_3 = u_{i'k'}u_{k'l'}u_{l'i'} = \eta_1^3 + \eta_2^3 + \eta_3^3, \tag{51}$$

where η_1, η_2, and η_3 are the principal values. Landau thus writes the following expansion of $\mathcal{E}_e$ to cubic order in $u_{i'k'}$ as follows:

$$\mathcal{E}_e = \mu L_2 + (\tfrac{1}{2}K - \tfrac{1}{3}\mu)L_1^2 + \tfrac{1}{3}AL_3 + BL_2L_1 + \tfrac{1}{3}CL_1^3, \tag{52}$$

where μ and K are referred to as the second-order elastic constants, and A, B, and C as the third-order elastic constants.

Equation (52), which is a general expansion for isotropic elastic media, may be simplified for incompressible media. Begin by writing the Jacobian as

$$J = \det(\partial x_i/\partial \tilde{x}_{i'}) = 1 + \partial u_i/\partial \tilde{x}_i + O[(\partial u/\partial \tilde{x})^2], \tag{53}$$

where the relation $x_i = \tilde{x}_i + u_i$ follows from the definition of particle displacement. Since $J = 1$ for incompressible media one concludes in this case that $\partial u_i/\partial \tilde{x}_i = O[(\partial u/\partial \tilde{x})^2]$. Therefore $L_1 = O(u_{i'k'}^2)$, and the terms in Eq. (52) containing L_1 are $O(u_{i'k'}^4)$ or smaller, such that up to cubic order in $u_{i'k'}$ one obtains

$$\mathcal{E}_e = \mu L_2 + \tfrac{1}{3}AL_3 \tag{54}$$

for incompressible media, which is consistent with the order of approximation in Eq. (52).

To evaluate P_e with Eq. (54) we calculate U_e using the integral in Eq. (10), which requires the invariants to be expressed as functions of $\tilde{r}$. The relation between the Lagrangian strain tensor and the Green deformation tensor is[6]

$$u_{i'k'} = \tfrac{1}{2}(G_{i'k'} - \delta_{i'k'}), \tag{55}$$

where $\delta_{i'k'}$ is the Kronecker delta, and therefore the principal values of these two tensors are related as

$$\eta_i = \tfrac{1}{2}(\nu_i^2 - 1). \tag{56}$$

From Eqs. (24)

$$\eta_1=\eta_2=\tfrac{1}{2}[(r/\tilde{r})^2-1], \quad \eta_3=\tfrac{1}{2}[(\tilde{r}/r)^4-1], \tag{57}$$

which permits the invariants in Eq. (54) to be expressed as powers of $r/\tilde{r}$.

The relation between r and $\tilde{r}$ is given by Eq. (27), which is rewritten here as

$$(r/\tilde{r})^n=[1+3x^3(\xi_0+\xi_0^2+\tfrac{1}{3}\xi_0^3)]^{n/3}, \tag{58}$$

where $x=R_0/\tilde{r}$ as in Eqs. (37), and

$$\xi_0=\delta R/R_0=(R-R_0)/R_0. \tag{59}$$

The integral in Eq. (10) can be evaluated analytically when Eq. (58) is expanded in powers of ξ_0. Up to cubic order in ξ_0 one obtains

$$(r/\tilde{r})^n=1+nx^3\xi_0+n[x^3+\tfrac{1}{2}(n-3)x^6]\xi_0^2 + n[\tfrac{1}{3}x^3+(n-3)x^6+\tfrac{1}{6}(n-3)(n-6)x^9]\xi_0^3. \tag{60}$$

Using Eq. (60) to expand Eqs. (50) and (51) to the same order of approximation and substituting the results in Eq. (54) yield

$$\mathcal{E}_e=6\mu(x^6\xi_0^2+2x^6\xi_0^3-5x^9\xi_0^3)-2Ax^9\xi_0^3. \tag{61}$$

Equation (10) can now be integrated to obtain

$$U_e=8\pi R_0^3\left[\mu\left(\frac{\delta R}{R_0}\right)^2-\frac{1}{2}\left(\mu+\frac{A}{3}\right)\left(\frac{\delta R}{R_0}\right)^3\right]. \tag{62}$$

Finally, substitution in Eq. (12), with the first-order expansion $R_0^2/R^2=1-2(\delta R/R_0)$, yields

$$P_e=4\mu\frac{\delta R}{R_0}-(11\mu+A)\left(\frac{\delta R}{R_0}\right)^2. \tag{63}$$

A comparison of Eqs. (63) and (47) indicates that at leading order the relation between the nonlinear moduli in the Mooney and Landau models for incompressible media is

$$A/\mu=2\beta-6. \tag{64}$$

Alternatively, one may write $\delta R/R_0=\zeta-1$ and differentiate Eq. (63) to obtain

$$P_e'(1)=4\mu, \quad P_e''(1)=-22\mu-2A, \tag{65}$$

and compare with Eqs. (45).

Figure 4 shows a comparison of Eq. (63) (dashed line) with the exact relation based on the Mooney potential, Eq. (39) (solid line), for $\beta=0$ and therefore $A/\mu=-6$ from Eq. (64). The two models are seen to be interchangeable for R within about 20% of R_0.

VI. RELATIVE INFLUENCE OF SHEAR STRESS

We now return to Eq. (11), rewritten here after division by R:

$$\ddot{R}+\frac{3}{2}\frac{\dot{R}^2}{R}=\frac{1}{\rho R}\left[P_g\left(\frac{R_1}{R}\right)^{3\gamma}-P_\infty-P_e(R)\right]. \tag{66}$$

An expansion of this equation is sought up to quadratic order in terms of the relative displacement about the equilibrium position R_1,

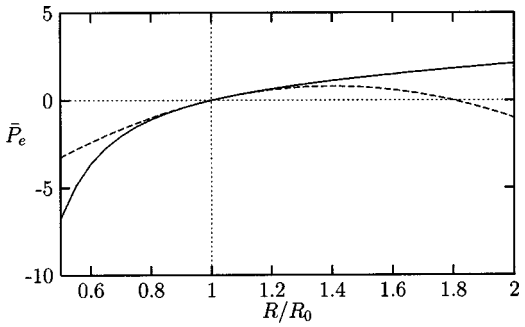


FIG. 4. Comparison of $\bar{P}_e=P_e/\mu$ calculated using Mooney's model, Eq. (39) with $\beta=0$ (solid line), with a calculation based on the quadratic expression corresponding to Landau's model, Eq. (63) with $A/\mu=-6$ (dashed line).

$$\xi=(R-R_1)/R_1, \tag{67}$$

where the equilibrium radius R_1 is determined by Eq. (13). For $R_1=R_0$, ξ reduces to ξ_0 in Eq. (59). Expansion of the gas pressure gives

$$P_g\left(\frac{R_1}{R}\right)^{3\gamma}=P_g\left[1-3\gamma\xi+\frac{3}{2}\gamma(3\gamma+1)\xi^2\right], \tag{68}$$

and Eq. (41) can be rewritten as

$$P_e(R)=P_g-P_\infty+R_1P_e'(R_1)\xi+\tfrac{1}{2}R_1^2P_e''(R_1)\xi^2, \tag{69}$$

where Eq. (13) was used to replace $P_e(R_1)$ by P_g-P_∞, and the primes here indicate derivatives with respect to R. Substitution in Eq. (66) together with $R=R_1(1+\xi)$ yields, up to quadratic order,

$$\ddot{\xi}+(\omega_g^2+\omega_e^2)\xi=(b_g+b_e)\xi^2-\tfrac{3}{2}\dot{\xi}^2, \tag{70}$$

where

$$\omega_g^2=\frac{3\gamma P_g}{\rho R_1^2}, \quad \omega_e^2=\frac{P_e'(R_1)}{\rho R_1}, \tag{71}$$

$$b_g=\frac{9\gamma(\gamma+1)P_g}{2\rho R_1^2}, \quad b_e=\frac{2P_e'(R_1)-R_1P_e''(R_1)}{2\rho R_1}. \tag{72}$$

Linear terms are collected on the left-hand side of Eq. (70), quadratic terms on the right. Coefficients with subscripts g and e identify terms that account for potential energy stored in the gas and the elastic medium, respectively, with the terms $\ddot{\xi}$ and $\dot{\xi}^2$ accounting for kinetic energy stored in the elastic medium. When $P_g=P_\infty$ and therefore $R_1=R_0$, Eq. (70) corresponds to Eq. (6) of Ostrovsky[8] for the perturbation in bubble volume when the Mooney potential is used to evaluate P_e' and P_e''.

In the linear approximation Eq. (70) reduces to

$$\ddot{\xi}+(\omega_g^2+\omega_e^2)\xi=0. \tag{73}$$

We begin by considering the effect of shear stress when $P_g=P_\infty$, in which case $R_1=R_0$ and, from either of Eqs. (45) and (65), $P_e'(R_0)=4\mu/R_0$. The natural frequency is seen to be $(\omega_g^2+\omega_e^2)^{1/2}$, where ω_g is the Minnaert frequency for oscillation of a gas bubble in an incompressible liquid,[12] and ω_e is the oscillation frequency of a spherical cavity in an

incompressible elastic medium in vacuum.[4] Equations (71) may be rewritten for $P_g = P_\infty$ as follows:

$$\omega_g^2 = 3\frac{\rho_g}{\rho}\frac{c_g^2}{R_0^2}, \qquad \omega_e^2 = 4\frac{c_t^2}{R_0^2}, \tag{74}$$

where $c_g = (\gamma P_g/\rho_g)^{1/2}$ is the sound speed in the gas, ρ_g is the density of the gas, and $c_t = (\mu/\rho)^{1/2}$ is the transverse (shear) wave speed in the elastic medium. The contributions due to potential energy stored in the gas and the elastic medium are comparable in the linear approximation when $\omega_g = \omega_e$, which occurs for $c_t = c_{t0}$, where

$$c_{t0}^2 = \frac{3\rho_g}{4\rho}c_g^2. \tag{75}$$

With $c_g \sim 340$ m/s and $\rho_g \sim 1.2$ kg/m^3 for air, and $\rho \sim 1000$ kg/m^3 for elastic media having densities similar to that of water, Eq. (75) yields $c_{t0} \sim 10$ m/s. For $c_t^2 \ll c_{t0}^2$ the shear stress may be ignored and the natural frequency is ω_g. For $c_t^2 \gg c_{t0}^2$ the potential energy in the gas may be ignored, and ω_e is the natural frequency.

We now reconsider the relative influence of shear stress on Eq. (73) for $P_g \neq P_\infty$, in which case $R_1 \neq R_0$. The shear wave speed corresponding to equal contributions due to potential energy in the gas and elastic medium is again determined by setting $\omega_g = \omega_e$, for which $R_1 P_e'(R_1) = 3\gamma P_g$ from Eqs. (71). Since $P_g = P_e(R_1) + P_\infty$ from Eq. (13), the condition $\omega_g = \omega_e$ is equivalent to

$$R_1 P_e'(R_1) = 3\gamma P_e(R_1) + 3\gamma P_\infty. \tag{76}$$

The effective pressure $P_e(R_1)$ is evaluated inside the elastic medium at the bubble wall, $R = R_1$, and P_∞ is the atmospheric pressure exerted on the outer surface of the medium at $R = \infty$. We choose the reference sound speed for the gas to be $c_g = (\gamma P_\infty/\rho_\infty)^{1/2}$, where ρ_∞ is the density of the gas at atmospheric pressure (the equilibrium density of the gas inside the bubble is different, corresponding to pressure P_g). The last term in Eq. (76) is now rewritten as follows: $3\gamma P_\infty = 3\rho_\infty c_g^2 = 4\rho c_{t0}^2 = 4\mu c_{t0}^2/c_{t1}^2$. Here, c_{t0} is again the value of c_t at which $\omega_g = \omega_e$ for $R_1 = R_0$ (i.e., for $P_g = P_\infty$), defined by Eq. (75) and taking into account that $\rho_g = \rho_\infty$ in this case, and c_{t1} is the corresponding value of c_t for $R_1 \neq R_0$. Substitution in Eq. (76) yields

$$c_{t1}^2 = \frac{4\mu c_{t0}^2}{R_1 P_e'(R_1) - 3\gamma P_e(R_1)}. \tag{77}$$

Since $P_e(R_0) = 0$ and $P_e'(R_0) = 4\mu/R_0$, Eq. (77) gives $c_{t1} = c_{t0}$ for $R_1 = R_0$, as required.

The ratio c_{t1}/c_{t0} is plotted in Fig. 5(a) as a function of R_1/R_0 for $\gamma = 1.4$, and with P_e and P_e' calculated for the Mooney potential according to Eqs. (39) and (43), respectively. The dependence on β is seen to be relatively weak. For $R_1/R_0 < 1$ ($P_g < P_\infty$) one obtains $c_{t1} < c_{t0}$ for any value of β. Reducing P_g below P_∞ thus reduces the value of μ at which shear stress becomes important. The reverse is true for $R_1/R_0 > 1$ ($P_g > P_\infty$), such that increasing P_g above P_∞ increases the value of μ at which shear stress becomes important. The result $c_{t1}/c_{t0} \to \infty$ for $R_1/R_0 \simeq 1.2$ indicates that for larger equilibrium bubble radii, associated with greater equilibrium gas pressures, the influence of shear stress is less than that of gas pressure for all values of μ.

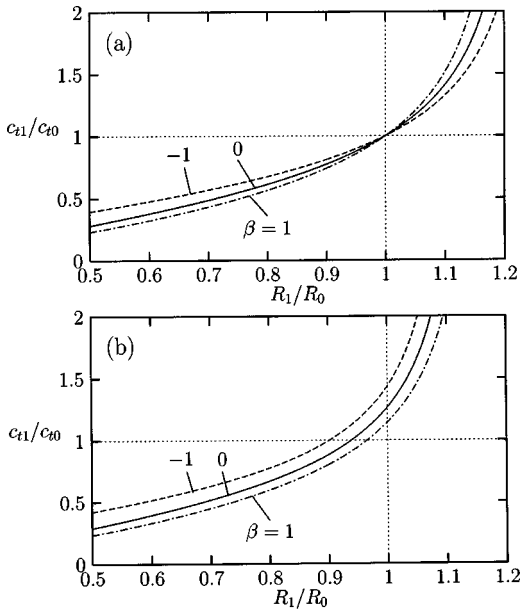


FIG. 5. The ratio c_{t1}/c_{t0} as a function of R_1/R_0 for $\gamma = 1.4$, with P_e and its derivatives calculated for the Mooney potential according to Eqs. (39), (43), and (44). (a) Calculations based on Eq. (77), corresponding to linear effects. (b) Calculations based on Eq. (78), corresponding to quadratic effects.

For completeness we consider also the relative effect of shear stress on the nonlinear terms in Eq. (70). This is accomplished by setting $b_g = b_e$ and proceeding as above to obtain

$$c_{t1}^2 = \frac{12\mu(\gamma+1)c_{t0}^2}{2R_1 P_e'(R_1) - R_1^2 P_e''(R_1) - 9\gamma(\gamma+1)P_e(R_1)}. \tag{78}$$

The ratio c_{t1}/c_{t0} based on Eq. (78) is plotted in Fig. 5(b) under the same conditions, and the dependence is seen to resemble that in Fig. 5(a), obtained in the linear approximation.

VII. CONCLUSION

The purpose of this paper is to assess the influence of shear stress in the medium surrounding a bubble, relative to compressive stress due to gas pressure within the bubble, on the dynamical behavior of the bubble. A Rayleigh–Plesset form of the equation describing the bubble dynamics is developed for the analysis. The investigation incorporates two alternative constitutive models for the elastic medium, due to Mooney and Landau, and it accounts for equilibrium bubble radii resulting from equilibrium gas pressures that may differ substantially from the pressure at infinity. The model can be modified to describe elastic shells and layers of finite thickness simply by dividing up the integration regions in the expressions for the energy functions. Numerical solutions of the dynamical equation for bubble radius as a function of time are not presented here. Meaningful simulations should take into account additional competing effects such as viscosity and radiation loss, which shall be considered in a future publication.

ACKNOWLEDGMENTS

SYE was supported by the U.S. Army Medical Research and Material Command. The other authors were supported by the Office of Naval Research, and by the Internal Research and Development program at Applied Research Laboratories, The University of Texas at Austin.

[1] J. S. Allen and R. A. Roy, ''Dynamics of gas bubbles in viscoelastic fluids. I. Linear viscoelasticity,'' J. Acoust. Soc. Am. **107**, 3167–3178 (2000).
[2] J. S. Allen and R. A. Roy, ''Dynamics of gas bubbles in viscoelastic fluids. II. Nonlinear viscoelasticity,'' J. Acoust. Soc. Am. **108**, 1640–1650 (2000).
[3] D. B. Khismatullin and A. Nadim, ''Radial oscillations of encapsulated microbubbles in viscoelastic liquids,'' Phys. Fluids **14**, 3534–3557 (2002).
[4] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, 3rd ed. (Pergamon, New York, 1986).
[5] E. Meyer, K. Brendel, and K. Tamm, ''Pulsation oscillations of cavities in rubber,'' J. Acoust. Soc. Am. **30**, 1116–1124 (1958).
[6] A. C. Eringen and E. S. Suhubi, *Elastodynamics* (Academic, New York, 1974), Vol. 1.
[7] V. N. Alekseev and S. A. Rybak, ''Gas bubble oscillations in elastic media,'' Acoust. Phys. **45**, 535–540 (1999).
[8] L. A. Ostrovsky, ''Nonlinear acoustics of slightly compressible porous media,'' Sov. Phys. Acoust. **34**, 523–526 (1988).
[9] K. Naugolnykh and L. Ostrovsky, *Nonlinear Wave Processes in Acoustics* (Cambridge University Press, New York, 1998).
[10] M. Mooney, ''A theory for large elastic deformation,'' J. Appl. Phys. **11**, 582–597 (1940).
[11] R. S. Rivlin, ''Large elastic deformations of isotropic materials. IV. Further developments of the general theory,'' Philos. Trans. R. Soc. London, Ser. A **241**, 379–397 (1948).
[12] T. G. Leighton, *The Acoustic Bubble* (Academic, New York, 1994).
[13] A. I. Lurie, *Nonlinear Theory of Elasticity* (Elsevier, New York, 1990).

Forward scattering from the sea surface and the van Cittert–Zernike theorem

Peter H. Dahl
Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105

(Received 8 June 2003; revised 17 October 2003; accepted 30 October 2003)

The van Cittert–Zernike theorem is used to generate models for the spatial coherence of a sound field that has been forward scattered from the sea surface. The theorem relates the spatial coherence of an observed wave field to the distant source intensity distribution associated with this field. In this case, the sea surface upon ensonification is taken to be the source, and the sea-surface bistatic cross section corrected for transmission loss is taken as a surrogate for the source intensity distribution. Improvements in methodology for generating an estimate of the 2D autocorrelation function for sea surface waveheight variation, necessary to compute the bistatic cross section, are documented in the Appendix. Upon invoking certain approximations, simple expressions for the characteristic length scales of vertical, horizontal, and horizontal–longitudinal coherence, are derived from the theorem. The three coherence length scales identify a coherence volume for the spatial coherence of a sound field arriving via the surface bounce channel. Models for spatial coherence derived from the van Cittert–Zernike theorem without these approximations compare reasonably well with measurements of complex vertical coherence made at 8 kHz and 20 kHz in the East China Sea as part of the 2001 ASIAEX field program. In terms of the ASIAEX field geometries and sea-surface conditions, at frequency of 20 kHz the coherence volume is a vertical layer 0.5 m thick by 3 m in each of the two horizontal dimensions; at 8 kHz these dimensions increase by a factor of 2.5, representing the ratio of the two frequencies. [DOI: 10.1121/1.1635417]

PACS numbers: 43.30.Hw, 43.30.Zk, 43.30.Re [SLB] Pages: 589–599

I. INTRODUCTION

Forward scattering from the sea surface is an important channel through which high-frequency underwater sound is transmitted and used for communication and imaging applications. Spatial coherence of the sound field received through this channel can be estimated by constructing probability density functions (PDF) that describe the spread in arrival angle at the receiver position.[1,2] For example, the PDFs $P_v(\theta_v)$ for vertical arrival angle and $P_h(\theta_h)$ for horizontal arrival angle are, to a very good approximation (provided that the angular spread is small), the Fourier transforms of vertical and horizontal spatial coherence, respectively. The approach from Refs. 1 and 2 is to construct the angular PDFs by summing scattered intensities associated with discrete vertical and horizontal arrival angles, normalizing the result into a PDF, then taking the Fourier transform. The sea-surface bistatic cross section σ modified by transmission loss to and from the sea surface is used for this purpose as a surrogate for the scattered intensity.

The Fourier transform duality between spatial coherence and angular spread of a received sound field is well known and utilized in underwater acoustics;[3,4] the conjugate Fourier variables in this relation are arrival angle and kd, where k is the wave number and d is the receiver separation on the array. From an operational standpoint, spatial coherence sets the effective aperture length of an array for directional discrimination.[3] From the standpoint of acoustical oceanography, spatial coherence, a readily observable quantity, yields information about the angular spreads caused by scattering from internal waves,[4] and other sources of granularity in the sound-speed field. In the case of forward scattering from the sea surface, the angular spread is determined by surface wave spectral properties, acoustic frequency, and geometry.[1,2]

Modeling spatial coherence through its Fourier transform relation with angular spread has considerable interpretive merit, but such an approach is cumbersome, particularly when spatial coherence along an arbitrary dimension is desired. Here, the van Cittert–Zernike theorem (vCZT) is applied to the problem of modeling spatial coherence associated with forward scattering from the sea surface; the vCZT both solves this problem and has inherent interpretive appeal. The vCZT is one of the most important theorems of statistical optics[5,6] and radio astronomy,[7] and among its many applications, it provides the mathematical foundation for astronomical interferometers to image celestial bodies using large arrays of radio telescopes. Briefly, in this application the vCZT relates the spatial coherence function of a wave field measured by an array of telescopes, to the Fourier transform of the brightness (or intensity) distribution for the celestial object emitting the radiation, provided certain approximations akin to a far field are invoked. An implicit assumption of the theorem is that the celestial object emits radiation that is both (approximately) incoherent and narrow band. Applying the vCZT to ocean acoustics, the sea surface is taken to be an incoherently radiating source upon activation by an incident sound field. The requirement for incoherent radiation generally means limiting to situations for which $kH \gg 1$, where H is the sea-surface rms waveheight; for natural sea surfaces this implies a nominal frequency range of O(10) kHz and above. The restriction on bandwidth[5,6] means that the path-length difference between a source point and

any two receiving points, Δs, must be such that $\Delta s/c \ll 1/\Delta f$, where c is nominal sound speed and Δf is the characteristic bandwidth of the sound field. That is, the channel through which the field propagates cannot temporally decorrelate within the time span of $\Delta s/c$. In the context of the ocean acoustic field geometries discussed here, $\Delta s/c \leqslant 0.1$ ms for the most energetic paths, and the characteristic bandwidth of the sound field will instead be set by pulse length.

Mallart and Fink[8] investigated theoretically and experimentally the vCZT applied to ultrasonic pulse echo measurements. They derived a formalism for their application based on scattering functions, and experimentally verified properties of backscattering from random media predicted by the vCZT when the same aperture was used to transmit and receive. The most fundamental property is that spatial covariance of the backscattered field originating from a uniform distribution of scatterers is proportional to the autocorrelation of the transmitter's aperture function. As mentioned above, the key assumption here is that a random scattering medium acts as an incoherent source for backscattering when ensonified, and it follows that the equivalent aperture of this incoherent source is set by that of the transmitter. Note that it is not essential to invoke the vCZT to derive this property of backscattering from uniform random media. Jackson and Moravan[9] arrived at this same result using an analytical point-scatterer model, and also verified with field measurements of sea-surface and ocean volume reverberation that the spatial covariance of such reverberation is reduced upon reducing the width of the transmitting array. Mallart and Fink[8] also investigated scattering from nonuniform media, which is handled naturally by the vCZT. They incorporated a medium outer-scale size into their scattering function, and verified experimentally that spatial covariance of the field backscattered from a nonuniform media is widened in the manner predicted by the theorem.

Section II describes how the vCZT is applied to problems in forward scattering from the sea surface, which is different from the case of backscattering from random media. The relation of the Fourier transform approximation to the vCZT is investigated and important scaling relations for the spatial coherence of the forward-scattered sound field are derived. In Sec. III previously published measurements of horizontal spatial coherence are revisited to demonstrate the scaling relation for horizontal coherence. In Sec. IV the vCZT, without resorting to the Fourier transform approximation, is used to model recent measurements of complex vertical coherence taken in the East China Sea.

II. THE VAN CITTERT–ZERNIKE THEOREM AND FORWARD SCATTERING FROM THE SEA SURFACE

The vCZT is applied to problems in high-frequency forward scattering from the sea surfaces as follows. An acoustic source S (Fig. 1) is at depth SD and a central receiver R_0 is at depth RD and range R. The (x,y,z) coordinate system is centered at R_0 and the coordinates of S relative to this system are $(-R,0,RD-SD)$. A second receiver R_1 is located on the y axis, and the pair of receivers might represent, for example, two elements of a horizontal line array aligned with this axis.

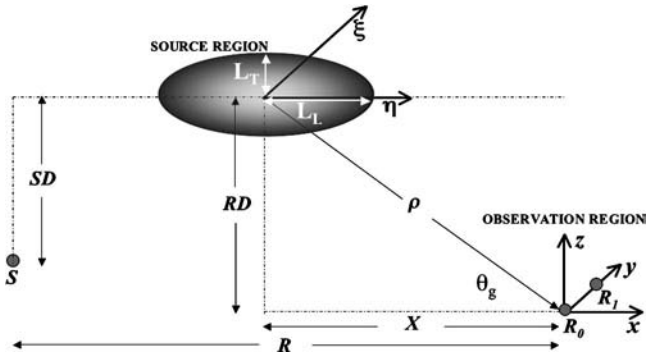


FIG. 1. Geometry for applying the van Cittert–Zernike theorem to problems in forward scattering from the sea surface showing source region described by $I(\tau,\eta)$, observation region, and key geometric variables.

In general, receivers are located anywhere within an observation region close to R_0; observation region is defined subsequently.

On the sea surface there exists an intensity distribution $I(\xi,\eta)$ equal to the acoustic intensity that has been forward scattered by the sea surface in directions towards the central receiver R_0 as a result of ensonification by the source at S. The two-dimensional ξ,η coordinate system is centered at position $(-X,0,RD)$ with respect to the x,y,z system, and because the surface is rough, the mean surface is taken to coincide with the ξ,η plane. The distance X is defined such that the origin of the ξ,η coordinate system coincides with the point corresponding to specular reflection from source S to receiver R_0, with X equal to $R\times RD/(SD+RD)$. The key parameter for characterizing forward scattering is the grazing angle θ_g associated with the nominal specular point.

For sea-surface forward scattering in the O(10)-kHz frequency range, $I(\xi,\eta)$ is effectively established by the bistatic cross-section distribution over the sea surface, $\sigma(\xi,\eta)$, which itself is a function of geometry (SD, RD, and R) and environment (sea-surface roughness correlation function). Measured sea-surface environmental data are used to compute $\sigma(\xi,\eta)$[1] and $\sigma(\xi,\eta)/r_{\rm inc}^2(\xi,\eta)$ is taken as a surrogate for $I(\xi,\eta)$, where $r_{\rm inc}$ is the slant range from S to location (ξ,η) on the sea surface. The effective domain of $I(\xi,\eta)$ will typically be long in the range direction and narrow in the transverse direction, given that omnidirectional sources and receivers are used in conjunction with sufficiently long pulse lengths. However, transmit and receive beam patterns, and pulse-length effects can truncate or otherwise shade $I(\xi,\eta)$; if required, a shading function can be applied in a straightforward manner to estimate spatial coherence via the vCZT.

The observation region is identified by being associated with one source region or $I(\xi,\eta)$. That is, for the purpose of computing spatial coherence via the vCZT, $\sigma(\xi,\eta)$ is computed for a given source depth, range, and one receiver location; the one corresponding to the nominal center of the receiving array at R_0. Nearby receivers subject to the coherence analysis and within a radius δ experience the same source intensity distribution on the sea surface. Numerical studies define an observation region by $\delta/\rho<0.1$, where ρ is the slant range from the specular point on the sea surface to

R_0. Note that δ is not the same as the coherence length scale discussed subsequently.

To investigate the relation of the Fourier transform approximation to the vCZT as applied to forward scattering from the sea surface, we use as a convenient artifice a Gaussian representation for the scattered intensity, $I(\xi,\eta) = \exp[-(\xi/L_T)^2-(\eta/L_L)^2]$. Here, the source region is now definitively bounded by the two length scales, L_L, for the longitudinal or range scale length, and L_T, for the transverse scale length. Spatial coherence between receiver R_0 and receiver R_1 located at $(0,y_1,0)$ is obtained from the vCZT without resorting to the Fourier transform approximation[5]

$$\Gamma_{01}=\frac{1}{\sqrt{I_0 I_1}}\int \frac{I(\xi,\eta)e^{ik(r_0-r_1)}d\xi\, d\eta}{r_0 r_1}, \tag{1}$$

where $I_{i=0,1}$ is given by

$$I_i=\frac{\int I(\xi,\eta)d\xi\, d\eta}{r_i^2}, \tag{2}$$

and $r_i=r_i(\xi,\eta)$ is slant range between a point ξ,η on the sea surface and the ith receiver. The Fourier transform approximation necessitates that the linear dimensions of the source and observation regions be small in comparison with r_i. Let us arbitrarily impose this on the source region by placing such limits on L_T and L_L, and on the observation region by limiting the maximum separation between receivers R_0 and R_1. Given such limits, we take the binomial approximation for r_1

$$r_1\approx\rho-\frac{X\eta}{2\rho}+\frac{y_1^2-2\xi y_1+\xi^2+\eta^2}{2\rho}, \tag{3}$$

with similar approximation for r_0, and therefore $r_0-r_1 = (-y_1^2/2\rho)-(\xi y_1/\rho)$. The exponent in Eq. (1) suggests spatial coherence will ultimately depend on the relative distance between receiver elements and not their absolute location; thus, take Δy for y_1. Upon replacing the slant ranges in the denominators of the integrands in Eqs. (1) and (2) by the constant ρ, Eq. (1) becomes, apart from a phase factor, a Fourier transform relating the source region's intensity distribution $I(\xi,\eta)$ to the spatial coherence measured in the observation region. Interestingly, only the binomial approximation for slant ranges has been applied, which is ordinarily associated with the Fresnel region. The result for horizontal spatial coherence along the y axis (call this Γ_H) is

$$\Gamma_H(\Delta y)=\exp(i\Omega)\exp\left[-\left(\frac{kL_T\Delta y}{2\rho}\right)^2\right]. \tag{4}$$

Here, the phase is set by $\Omega=-k\Delta y^2/2\rho$, and can be ignored provided $k\Delta y^2\ll 2\rho$. In any case a coherence length scale d^* is identified by setting $|\Gamma_H|=e^{-1/2}$, which puts $d^* = 2\rho/kL_T$. That d^* increases with distance ρ from the source region is a fundamental result of the vCZT and its most important manifestation. With forward scattering from the sea surface, however, the source region is intimately linked to the overall geometry, and this manifestation takes a different form that we show subsequently. Continuing with the present example, spatial coherence along the other two axes is derived similarly. The result (evaluated as in the above to first power in quantity $1/\rho$) for vertical coherence $\Gamma_V(\Delta z)$ is

$$\Gamma_V(\Delta z)=\exp(i\Omega+ik\Delta z\sin\theta_g) \times\exp\left[-\left(\frac{kL_L\sin\theta_g\cos\theta_g\Delta z}{2\rho}\right)^2\right], \tag{5}$$

where $\Omega=[-k(\Delta z\cos\theta_g)^2]/2\rho$, and for horizontal–longitudinal coherence $\Gamma_{HL}(\Delta x)$ is

$$\Gamma_{HL}(\Delta x)=\exp(i\Omega-ik\Delta x\cos\theta_g) \times\exp\left[-\left(\frac{kL_L\sin^2\theta_g\Delta x}{2\rho}\right)^2\right], \tag{6}$$

where $\Omega=[-k(\Delta x\sin\theta_g)^2]/2\rho$. Likewise, the Ω component of phase can be ignored given conditions on k and ρ, but for vertical and horizontal–longitudinal coherence there exists an intrinsic phase that is a function of the nominal sea surface grazing angle θ_g.

The Gaussian representation for $I(\xi,\eta)$ is a useful artifice leading to an analytic result for the Fourier transform. But, to progress towards evaluating d^* for spatial coherence in high-frequency forward scattering from the sea surface requires finding an expression that embodies both the geometry and environmental conditions that determine the scattering. In doing so, we take from the preceding result that L_T/ρ is equivalent to a characteristic horizontal angular width for the scattered intensity originating from the source region. By hypothesis, assume this angular width is set by the high-level scattering occurring in the vicinity of the specular point. Here, scattering functional behavior can be represented by the quasispecular expression for the scattering cross section, which itself takes a Gaussian form whose width is established by the waveslope variance.[10,11] Using the bistatic form of this cross section[12] as in Eq. (9) of Ref. 14 (the sonar version of that in Ref. 12), it is readily shown[13] that in the vicinity of the specular point, the Gaussian is controlled by the bistatic angle ϕ_b and a $e^{-1/2}$ width is reached when $\phi_b \approx 2S_L\tan\theta_g$, where S_L^2 is the mean-square large-scale slope of the sea surface.[2] The ratio of ϕ_b to horizontal arrival angle θ_H is $\cos\theta_g/(1+RD/SD)$,[13] and thus $\sigma_{\theta_H} \sim S_L\sin\theta_g/(1+RD/SD)$, where σ_{θ_H} is defined as the rms horizontal angular spread. Using σ_{θ_H} in place of L_T/ρ gives

$$kd_H^*\sim\frac{(1+RD/SD)}{S_L\sin\theta_g}. \tag{7}$$

Equation (7) is derived in Ref. 13 without relating to the vCZT, and a multiplying factor of O(1) is employed there to fit the expression to field measurements. For vertical coherence along the z axis and horizontal–longitudinal coherence along the x axis, we take that $L_T/L_L\sim\sin^2\theta_g$ for high-frequency, incoherent forward scattering from the sea surface,[14] which leads immediately to

$$kd_V^*\sim\frac{(1+RD/SD)}{S_L\cos\theta_g}, \tag{8}$$

and

$$kd_{HL}^{*} \sim \frac{(1+RD/SD)}{S_L \sin\theta_g}, \tag{9}$$

(or the same as kd_H^*). Thus, in the observation region about R_0 in Fig. 1, we anticipate a *coherence volume* of approximate size $(2d_V^*)(2d_H^*)(2d_{HL}^*)$, the factor of 2 arising owing to the symmetry of coherence magnitude about the point of zero-spatial separation. One can imagine applications in underwater acoustic imaging and communication where this coherence volume can be exploited.

The preceding expressions for d_H^*, d_V^*, and d_{HL}^* emerging from the above approximations serve as only rough guidelines. They are useful for predicting changes in coherence given a change in geometry; for example, horizontal coherence evidently scales with the geometric parameter $g^*=(1+RD/SD)/\sin\theta_g$. Yet Eqs. (7) and (8) also predict the ratio of horizontal to vertical coherence length as equal to $\cos\theta_g/\sin\theta_g$, when in fact under certain conditions this ratio is substantially different. Such conditions originate in the approximation for r_0-r_1, where phase terms that are nonlinear in the ξ and η coordinates can no longer be ignored for low grazing angles. In particular, the nonlinear phase terms can only be ignored for the situations that satisfy $(\cos^2\theta_g)/[\sin^4\theta_g(kd_H^*)^2] \ll 1$, which in practice often requires that $\theta_g \geqslant 10°$.

III. COMPARING SCALING RELATION FOR HORIZONTAL COHERENCE TO FIELD DATA

The scaling relation for horizontal coherence in sea-surface forward scattering is compared here to field data from two experiments that measured horizontal spatial coherence at 30 kHz. One set was obtained at a range of 714 m in pelagic waters during which the 10-min averaged wind speed was 5 m/s. The acoustic source was suspended from a tethered spar buoy and the horizontal receiving array was mounted on the hull of the *R/P* FLIP at depth 66 m.[13] The other set was obtained at a range of 72 m in littoral waters during which the 10-min averaged wind speed was 4 m/s. The same acoustic source system and horizontal receiving array were used, but in this case the array was suspended from a research vessel at a depth of 9.5 m.[1] Each set consists of a pair of measurements taken sequentially within minutes of each other but between which the source depth was changed. For the 714-m measurements the source depth changed from 147 to 57 m, changing the nominal sea-surface grazing angle, θ_g, from 17° to 10°. For the 72-m measurements the source depth changed from 17.2 to 7.7 m, changing θ_g from 20° to 14°. The two sets of measurements each isolated the role of geometry while the sea-surface environment was held constant. This role, for horizontal coherence, is encapsulated by the parameter g^*, which emerged from the foregoing analysis of the Fourier transform approximation for the vCZT. Data from measurement pairs within either set, taken under exactly the same environmental conditions, are expected to collapse to a single line when scaled by this parameter, provided that the grazing angle criterion is satisfied. In terms of the latter, the parameter $(\cos^2\theta_g)/[\sin^4\theta_g(kd_H^*)^2]$ is 0.1–0.2 for these measurements, and thus the scaling relations should apply.

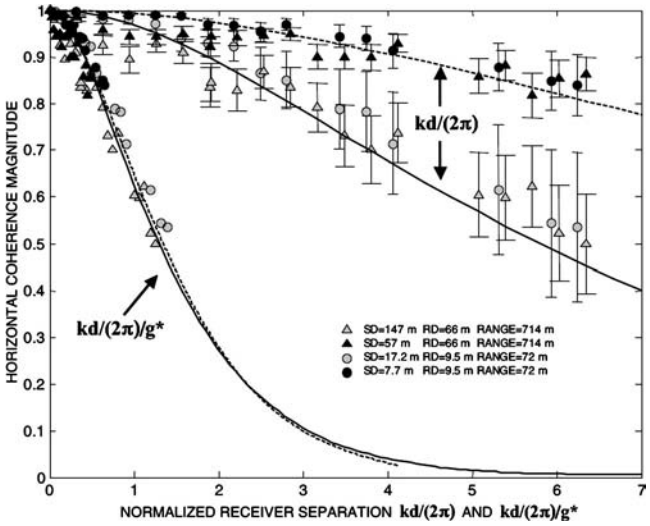


FIG. 2. Estimates of horizontal spatial coherence for a 30-kHz sound field forward scattered from the sea surface made at ranges 714 m (triangles) and 72 m (circles), using two different source depths at each range. The model curves are predictions for the 72-m range measurements; dashed line corresponds to 7.7-m source and solid line corresponds to 17.2-m source depth. Both data and model curves are plotted against receiver separation (d) normalized by wavelength, $kd/(2\pi)$, and against scaled separation, $kd/(2\pi)/g^*$ (see the text for values of g^*). Error bars for data points are shown only for the $kd/(2\pi)$ plotting.

This comparison alone serves to illustrate the horizontal scaling law. However, both sets of measurements also share some degree of similarity in the sea-surface environment, insofar as the 10-min averaged wind speeds are comparable. Similarity is necessarily confined to a wind-speed-dependent rms large-scale slope (see Ref. 2 for discussion of this parameter) because the rms waveheight for the swell-dominated pelagic conditions was 0.5 m compared to 0.07 m for the littoral measurements. Furthermore, the geometric scales are similar, i.e., the deep and shallow source depths for the 714-m range measurements put g^* equal to 5.0 and 12.4, respectively; for the 72-m range measurements, the deep and shallow source depths put g^* equal to 4.5 and 9.2, respectively. Therefore, deep and shallow source measurements in one set will be similar to their counterparts in the other set owing to similarity in both the environment and scaling parameter. This is shown in Fig. 2, for which the data from both sets are plotted first against receiver separation normalized by wavelength, $kd/(2\pi)$, and then against scaled receiver separation, $kd/(2\pi)/g^*$. For the nonscaled plotting the deep and shallow source measurements made at range 72 m roughly match their counterparts made at a factor of 10 greater in range. This result is quite apart from the fundamental result of the vCZT in Sec. II that has coherence length scale increasing linearly with range from the source, but is expected on the basis of g^*. Regarding the scaled plotting, it is clear that for both sets of measurements scaling by the parameter g^* effectively collapses the shallow and deep source measurements into one line. Furthermore, the four measurements from both sets nearly collapse into a single cluster of data, consistent with the similarity displayed in the unscaled plot. The model curves apply specifically to the 72-m range measurements; these were originally derived[1]

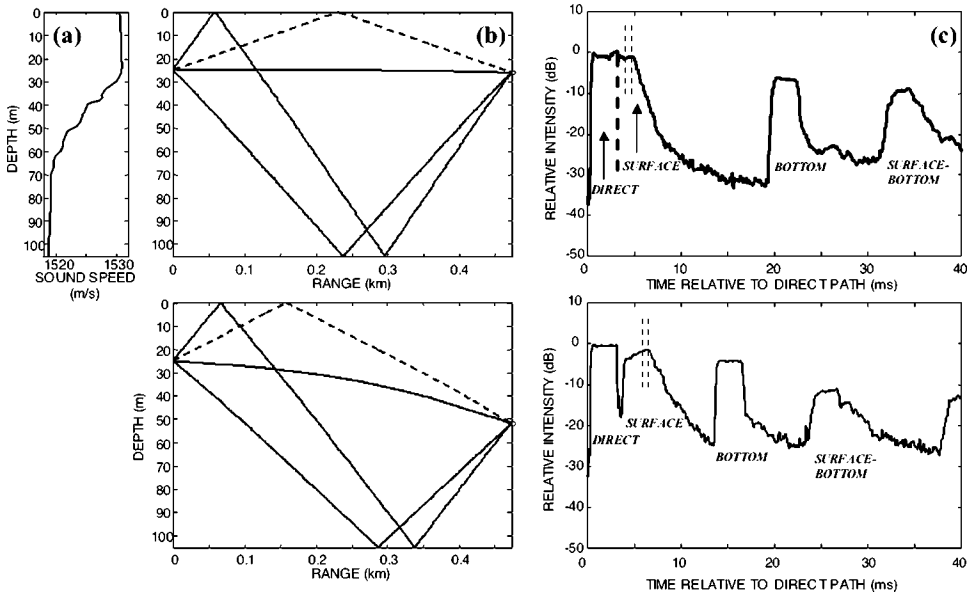


FIG. 3. (a) Average sound-speed profile corresponding to time of acoustic measurement sets 6–8. (b) Ray diagrams for 26-m depth VLA (top) and 52-m depth VLA (bottom) generated by sound-speed profile in (a) plot with dashed line corresponding to the nominal ray paths interacting once with the sea surface. (c) Intensity in dB relative to maximum corresponding to one element of the 26-m depth VLA (top) and of the 52-m depth VLA (bottom). Data are from set 7, frequency 20 kHz, and represent the average of 20 ping receptions. Vertical lines delineate the regions from which estimates of complex vertical coherence across each array are made.

by the approach involving a PDF for horizontal arrival angle; and deriving them using the approach involving the vCZT as in Sec. IV B yields an identical result. The two model curves very nearly collapse to one curve when plotted against $kd/(2\pi)/g^*$, using $g^*=4.5$ and 9.2, indicating the degree to which scaling ought to collapse horizontal spatial coherence data for this combination of geometry and sea-surface conditions.

IV. MEASUREMENTS OF COMPLEX VERTICAL COHERENCE FROM ASIAEX EAST CHINA SEA COMPARED TO MODELS OBTAINED FROM THE VAN CITTERT–ZERNIKE THEOREM

A. ASIAEX measurements

Measurements of complex vertical coherence were made in the East China Sea in May and June 2001 as part of the Asian Seas International Acoustics Experiment (ASIAEX) field program. A collection of papers covering results from all aspects of the ASIAEX field program carried out in both the South and East China seas is expected in coming months, and thus only details relevant to the particular measurements discussed here are given. The spatial coherence measurements were made by the Applied Physics Laboratory from the R/V MELVILLE operating 350 km east of Shanghai off the Chinese continental margin (29.65°N, 126.82°E) in waters 105 m deep. Acoustic signal receptions were recorded with two colocated vertical line arrays (VLA) each of length 1 m, with one array at depth 26 m and the other array at depth 52 m, referenced to the array center. Each VLA consisted of four elements (omnidirectional ITC 1042) with top-to-bottom element separation equal to 13, 30, and 60 cm. The two-array system was deployed at a nominal range of 500 m from the source; precise range depended on measurement time of day and was established acoustically. Data sampling rate was 50 kHz and data were sent back to the MELVILLE through an rf modem. The acoustic source was deployed off the stern of the MELVILLE while it maintained dynamic positioning, and the transmitting transducer (depth 25 m) was an omnidirectional ITC 1007. A family of seven transmit waveforms with center frequencies spanning 2–20 kHz was cycled every 100 s for 20 times; data corresponding to a set of 20×7 transmissions were identified as an experimental set, with this process repeated after a few seconds of delay. The vertical spatial coherence measurements presented here are limited to those obtained using transmit waveforms consisting of cw pulses of length 3 ms and center frequency equal to 8 and 20 kHz. Furthermore, data from three sequential sets of 20 receptions each (sets 6–8) are averaged to reduce statistical uncertainty in estimates of spatial coherence. Thus, the sea-surface environment is taken to be unchanging in terms of mean wind speed and wave spectral properties over approximately 6000 s.

The sea-surface environment was characterized by wind speed, measured with MELVILLE's IMET station, and the directional wave spectra, measured using a 0.9-m diameter TRIAXYS directional wave buoy deployed approximately 0.5 km from the MELVILLE's position. The buoy measured waveheight variance spectra in 0.005-Hz bins from 0.03 to 0.64 Hz and in 3-deg directional bins, with spectra estimated every 0.5 h based on a 20-min averaging time. Within the time window of the acoustic measurements, approximately 0630–0800 UTC 31 May 2001, average wind speed was $7\ \text{m/s}\pm0.5\ \text{m/s}$, and the rms sea-surface waveheight was $0.3\ \text{m}\pm0.01\ \text{m}$. Wind and waves originated from the north, and the acoustic line-of-site axis was aligned north-to-south. Further details concerning sea-surface wave spectral and directional properties are presented in the Appendix.

The sound-speed profile was monitored with CTD casts,

and Fig. 3(a) shows the averaged sound-speed profile for the same period of the acoustic measurements, ray diagrams (b) computed from the averaged profile, and corresponding multipath arrival structure (c) from averaged data recorded by a single element in the upper and lower vertical array. The focus here is the surface bounce path, represented by the dashed line in the ray diagrams. With the 3-ms pulse the surface bounce path is completely resolvable from the direct path for measurements made with the lower VLA, while there is partial overlap between these two paths for measurements made with the upper VLA. Vertical lines in the two center plots delineate regions where estimates of the complex envelope are made from each ping reception (after leading edge alignment), with extra care taken in the overlap case to ensure that received sound field associated only with the surface-interacting channel is sampled. An average value of the complex envelope within the vertical lines is computed for each ping reception recorded by the four receivers composing the upper and lower receiving arrays. Define p_i as the average complex envelope amplitude for the ith reception measured by receiver p, with the same for receiver q; then, an estimate of complex coherence $\hat{\Gamma}_{pq}$ is obtained via

$$\hat{\Gamma}_{pq}=\frac{\langle p_i q_i^* \rangle}{\sqrt{\langle p_i p_i^* \rangle \langle q_i q_i^* \rangle}}, \tag{10}$$

where brackets denotes an average over 60 ping receptions. With four receivers there are six combinations of relative receiver separation beyond zero separation. A good approximation for the standard deviation of the coherence magnitude estimate $|\hat{\Gamma}_{pq}|$ is $(1-|\hat{\Gamma}_{pq}|^2)/\sqrt{n}$, with n being the number of ping receptions entering the average.[15] From bootstrap simulation,[16] we conclude that this same expression, multiplied by $\sqrt{2}$, can be used to approximate the standard deviation of estimates for the *real* and *imaginary* parts of $\hat{\Gamma}_{pq}$.

B. Modeling and model/data comparison

The bistatic cross-section distribution over the sea surface, $\sigma(\xi,\eta)$, is computed using Eqs. (1), (5), and (6) from Ref. 1 that rely on the small slope approximation. The $\sigma(\xi,\eta)$ function depends explicitly on the incident, scattered, and bistatic angles (see Ref. 1), and these are determined implicitly by the scattering geometry, i.e., the source depth, receiver depth, range, and location on the sea surface. The sea-surface environmental dependence entering into $\sigma(\xi,\eta)$ resides in the two-dimensional autocovariance function for surface relief. The Appendix describes how this function is generated from a combination of wave buoy measurements and wave spectral modeling.

With a wind speed of 7 m/s, acoustically significant concentrations of near-surface bubbles are expected. Although such bubbles tend to dominate high-frequency, low-angle backscattering from the sea surface,[17] they have, at this wind speed, little or no measurable impact on spatial coherence in forward scattering from the sea surface.[2] For completeness, however, calculations of $\sigma(\xi,\eta)$ include a component due to bubbles [Eq. (13) of Ref. 1] based on the parameter β_1 equal to the area-normalized, depth-integrated extinction cross section. Equation (5) of Ref. 18 is used to estimate this parameter, which puts β_1 at 0.009 for 20 kHz and at 0.004 for 8 kHz. Figure 4 shows a plan view of $\sigma(\xi,\eta)$ corresponding to (a) the upper VLA geometry computed for a receiver at depth 26 m, and (b) the lower VLA geometry computed for a receiver at depth 52 m. In both cases the source depth is 25 m and the range is 467 m, and to reiterate the point made in Sec. II, each $\sigma(\xi,\eta)$ applies to all four receivers on their respective 1-m VLAs. Note: in Fig. 4, cross range corresponds to ξ and down range corresponds to η; otherwise, the $\sigma(\xi,\eta)$ distributions are not displayed according to a ξ,η coordinate system *per se* as defined in Sec. II.

Models for vertical spatial coherence corresponding to the upper and lower ASIAEX receiving arrays are derived by numerically implementing the van Cittert–Zernike theorem (vCZT) as expressed in Eq. (1). This means using the two distributions for bistatic cross section over the sea surface shown in Fig. 4 with each modified by one-way spreading from source to sea surface, $r_{\mathrm{inc}}^2(\xi,\eta)$, to obtain a surrogate for the intensity distribution $I(\xi,\eta)$. Additionally, no approximations are made for r_{inc}, and for the ranges r_0 and r_1 as was done in Sec. II that lead to a Fourier transform realization of Eq. (1). With minor changes in this numerical implementation, models for horizontal and horizontal–longitudinal coherence are derived for reference comparison. For a given acquisition geometry as defined by source depth, receiver depth, and range, one $I(\xi,\eta)$ function is common to all implementations of Eq. (1). There is no shading function applied to $I(\xi,\eta)$ because both source and receive elements used for ASIAEX field measurements were effectively omnidirectional.

Measurements of vertical coherence ($\hat{\Gamma}_V$) made with the upper and lower receiving arrays at 8 and 20 kHz compared with their model equivalents (Γ_V) are shown in Fig. 5. To simplify viewing, only the absolute value (upper plots) and real part (lower plots) of complex $\hat{\Gamma}_V$ and Γ_V are displayed. Error bars for $|\hat{\Gamma}_V|$ and $\mathrm{Re}(\hat{\Gamma}_V)$ represent $\pm$ the standard deviation of an estimate as defined above. Although the coherence measurements extend only to the maximum array length (1 m), model curves are plotted out to 2 m to illustrate trends. Furthermore, the contrast between measurements made at 8 and 20 kHz is improved by plotting against receiver separation rather than against separation normalized by wavelength, or $kd/(2\pi)$. Were the data plotted against $kd/(2\pi)$, measurements made at the two frequencies would nearly, but not completely, collapse to a single curve. This is because, for a given spatial separation in units of wavelength, spatial coherence at 8 kHz is slightly higher than at 20 kHz owing entirely to frequency-dependent differences in $\sigma(\xi,\eta)$. The data are consistent with this subtle trend, but given the measurement variability the trend is best illustrated by the model curves. The two vertical dotted lines in each of the upper plots identify the spatial separation at which $|\Gamma_V|$ reaches $e^{-1/2}$, and in terms of the number of wavelengths, the separation at 20 kHz is about 10% less than at 8 kHz.

The models for complex vertical coherence compare reasonably well with ASIAEX field measurements in both magnitude and real part. The models are initially computed

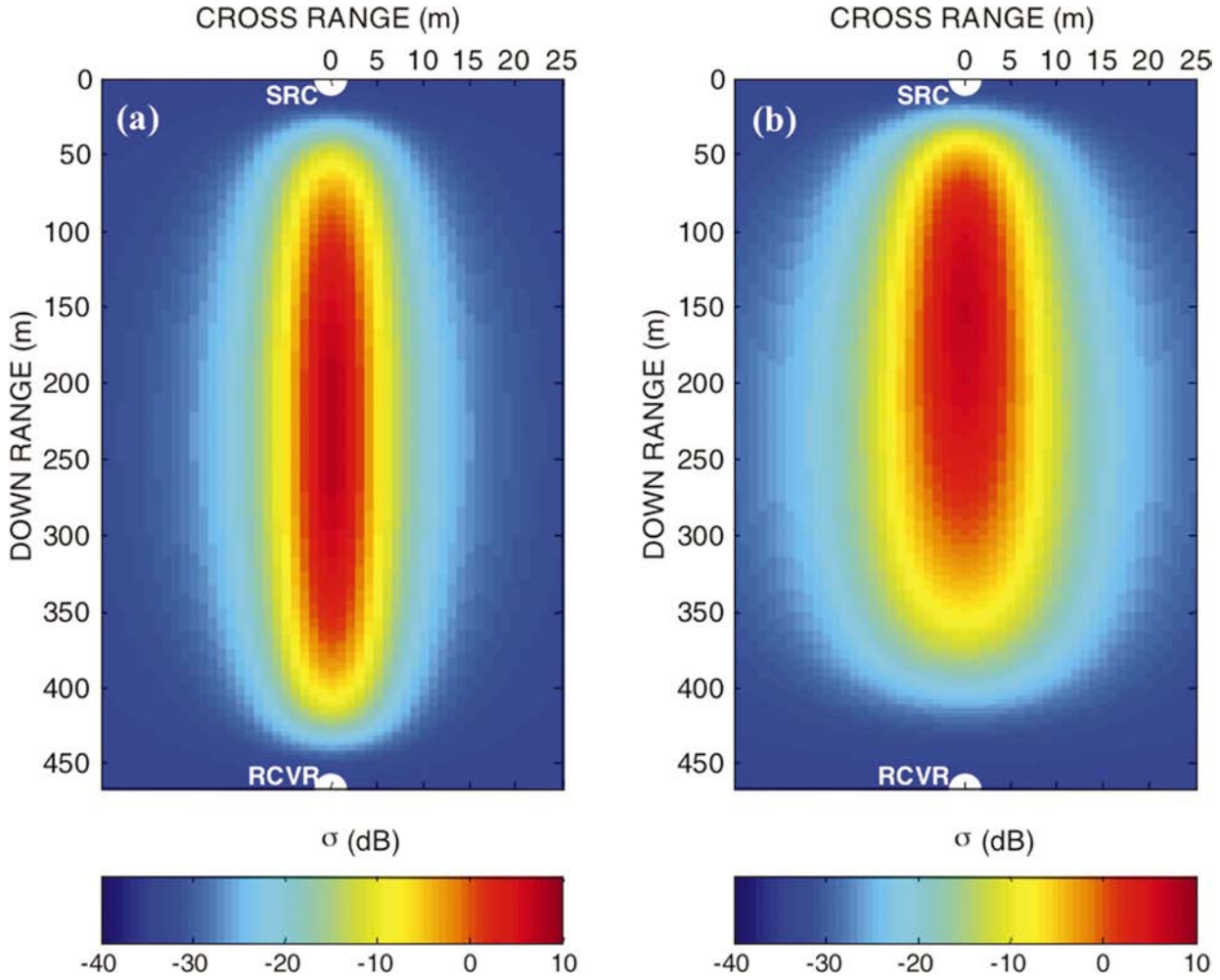


FIG. 4. Model estimates of the 20-kHz bistatic cross section σ expressed in dB, as function of position on the sea surface. Experimental conditions and geometry correspond to measurement sets 6–8: source depth 25 m, range 467 m; (a) receiver depth 26 m; and (b) receiver depth 52 m. Source and receiver positions are marked to orient the plan view; note that the aspect ratio exaggerates the cross-range dimension.

for iso-speed conditions, and as the primary influence of refraction on vertical spatial coherence is to alter the phase, effects due to refraction are incorporated into model results as follows. Based on the sound-speed profile shown in Fig. 3, the vertical arrival angles (call them θ_V) associated with the specular reflection paths for the upper and lower array are 6.1° and 10.9°, respectively; whereas in the absence of refraction these angles (call them θ_{Vo}) are 6.1° (i.e., unchanged at this precision) and 9.4°, respectively. The model results are corrected for refraction by changing the phase, that is $\Gamma'_V = \Gamma_V e^{-ik\Delta z(\sin\theta_{Vo} - \sin\theta_V)}$, where Γ'_V incorporates the phase angle change due to refraction. Solid lines in the lower plots of Fig. 5 show phase-corrected model curves for the

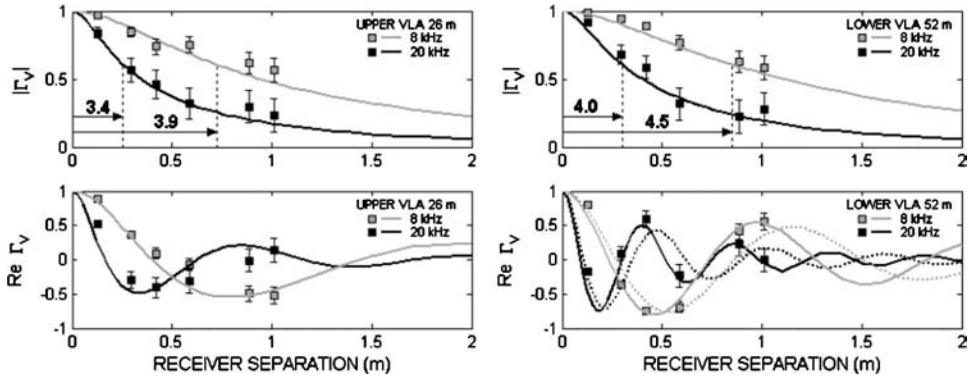


FIG. 5. Comparison of measured values of complex vertical coherence, in terms of absolute value (upper plots) and real part (lower plots), with model curves derived from the van Cittert–Zernike theorem. The number of wavelengths for which model curves for vertical coherence at 8 and 20 kHz reach the $e^{-1/2}$ level are listed above arrows in the upper plots. Dotted curves in the lower right plot are model representations for real part of vertical coherence under iso-velocity conditions.

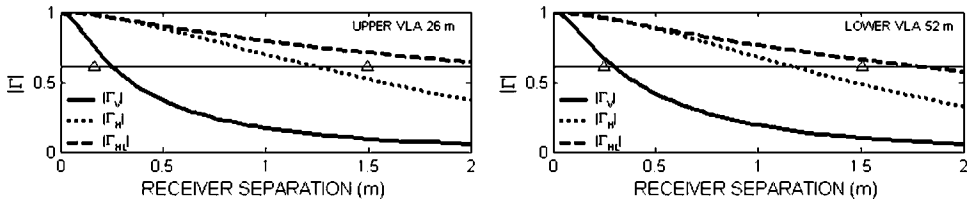


FIG. 6. Models for $|\Gamma_V|$ at 20 kHz from Fig. 5 compared with the corresponding model curves for $|\Gamma_H|$ and $|\Gamma_{HL}|$. Coherence length scales d_V^*, d_H^*, and d_{HL}^* are identified by the intersection of the horizontal line (representing a coherence criterion of $e^{-1/2}$) with the three model curves in each plot. The triangles in each plot represent estimates of these values based on the scaling relations derived in Sec. II (only two triangles shown as scaling relations put d_H^* equal to d_{HL}^*).

real part of vertical coherence, $\mathrm{Re}(\Gamma_V')$, and dashed lines (in the lower right plot only) show uncorrected curves corresponding to iso-speed conditions, $\mathrm{Re}(\Gamma_V)$, with the difference between the two discernible only in the lower-array data that was subjected to more refraction. In this case, the phase change clearly improves comparison between model curves and observations for the real part of complex vertical coherence.

C. Coherence volume and scaling relations

Figure 6 compares the models for $|\Gamma_V|$ at 20 kHz from Fig. 5 with the corresponding $|\Gamma_H|$ and $|\Gamma_{HL}|$ models, for which there are no field observations. Coherence length scales d_V^*, d_H^*, and d_{HL}^* are identified by the intersection of the dotted, horizontal line (representing a coherence criterion of $e^{-1/2}$) with the three model curves in each plot. A different coherence criterion can be used depending on application. In any case, in theory there exists a coherence volume associated with each receive depth of approximate size $(2d_V^*)(2d_H^*)(2d_{HL}^*)$ m^3; for example, for the 52-m receiver $(2d_V^*)\approx 0.5$ m, $(2d_H^*)\approx 2.5$ m, and $(2d_{HL}^*)\approx 3.5$ m. The coherence volume applies to a narrow-band acoustic field arriving via the sea-surface bounce path, or channel, from a 20-kHz source at depth 25 m and range 497 m. For an 8-kHz source, each dimension of the coherence volume should be increased by a factor of approximately 2.5, or the ratio of frequencies.

Using the d_H^* from the model curves, the parameter $(\cos^2\theta_g)/[\sin^4\theta_g(kd_H^*)^2]$ from Sec. II equals 0.11 for the receiver at 52 m but increases to 0.75 for the receiver at 26 m owing to its smaller grazing angle. Thus, we anticipate that the Fourier transform scaling relations derived in Sec. II will be less reliable, particularly for the shallower receiver. Nevertheless, these relations still yield, with overwhelming simplicity, a rough approximation for the length scales d_V^*, d_H^*, and d_{HL}^*. Based on a rms large-scale slope $S_L\sim 0.15$ (see the Appendix), values for d_V^* and d_H^* estimated from the scaling relations are shown by the triangles in Fig. 6. (Recall the scaling relations put d_H^* equal to d_{HL}^*.) But, it is also evident from Fig. 6 that $|\Gamma_H|$ and $|\Gamma_{HL}|$ are not the same, contrary to that predicted by the approximations invoked in Sec. II; yet, for spatial lags $\leqslant 0.5$ m these functions are more or less equivalent and could have identical coherence length scales established under a higher coherence criterion, say $e^{-1/4}$. Still, d_H^* from the scaling relations tends to roughly split the difference between the d_H^* and d_{HL}^* values associated with model curves for $|\Gamma_H|$ and $|\Gamma_{HL}|$ that are based on the $e^{-1/2}$ criterion. The scaling relations give estimates of d_V^* that are reasonably consistent with values associated with the model curves for $|\Gamma_V|$, with better agreement for the 52-m receiver consistent with its larger sea-surface grazing angle.

V. SUMMARY

The vCZT is used to generate models for the spatial coherence in a sound field forward scattered once from the sea surface and received on a vertical line array. The models compared reasonably well with measurements of complex vertical coherence made at 8 and 20 kHz in the East China Sea, as part of the ASIAEX field program. The vCZT is used here in conjunction with models for sea-surface bistatic cross section, and improvements in methodology for generating the necessary estimate of the two-dimensional autocorrelation function for sea-surface waveheight variation are documented in the Appendix.

The vCZT operates on a source intensity distribution; here, the source distribution is the sea surface upon ensonification by a narrow-band source. It is further assumed that the source is incoherent, which implies that $kH\gg 1$, where k is acoustic wave number and H is rms waveheight. Upon invoking approximations that depend on both array acquisition geometry and length scales of source distribution, the vCZT leads to a Fourier transform relation between spatial coherence and the source intensity distribution. Importantly, these approximations are not necessary to apply the vCZT, and, in the case of forward scattering from the sea surface, the approximations are in many situations not sustainable. Nevertheless, the Fourier transform relation and attendant approximations (akin to the Fresnel approximation) give way to important scaling relations. These are $kd_H^*, kd_{HL}^* \sim (1+RD/SD)/(S_L\sin\theta_g)$ and $kd_V^*\sim(1+RD/SD)/(S_L\cos\theta_g)$, where d_H^* is the characteristic scale for horizontal coherence (a direction perpendicular to the source-to-receiver direction), d_{HL}^* is the characteristic scale for horizontal–longitudinal coherence (a direction in line with the source-to-receiver direction), and d_V^* is the characteristic scale for vertical coherence. The scaling relations depend on the source depth (SD), receiver depth (RD), sea surface grazing angle (θ_g), and on an environmental parameter describing the effective long-wave slope of the sea surface (S_L).

The scaling relations for horizontal coherence were compared with data originating from two sets of archival horizontal coherence measurements made at 30 kHz. Each set consisted of a pair of measurements taken sequentially but between which the source depth was changed, with sub-

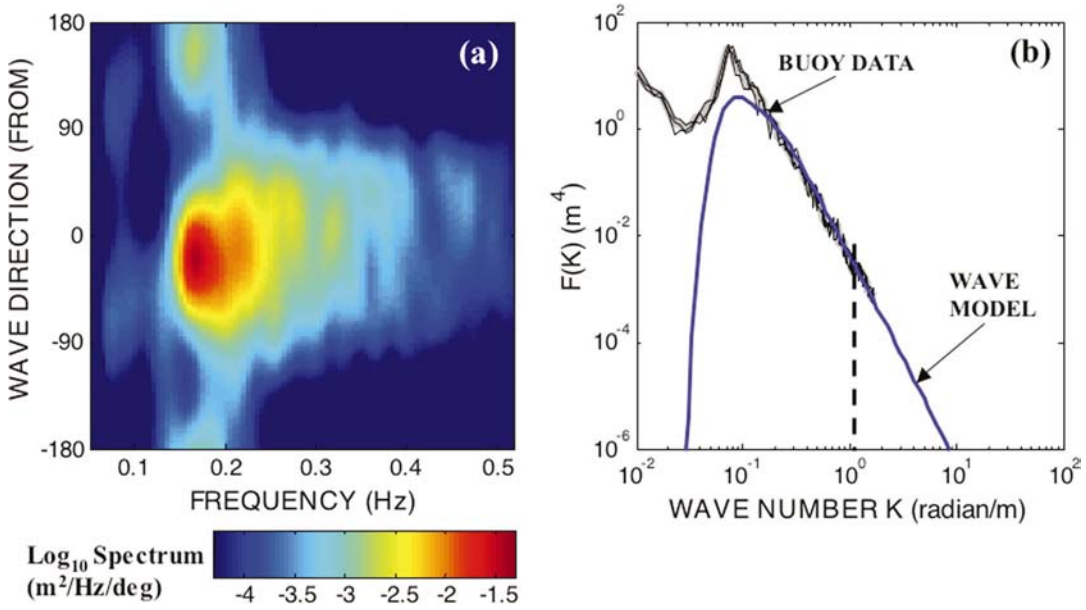


FIG. 7. (a) The directional frequency spectrum measured contemporaneously with the acoustic measurements, based on the average of three directional frequency spectra obtained during the period 0630–0800 UTC 31 May 2001. (b) Same three spectra each expressed as a directionally averaged wave number spectrum, or $F(K)$ (thin, black lines), their average (thick, gray line), and model $F(K)$ spectrum (Ref. 18) computed for a wind speed of 7 m/s and for fully developed conditions (blue line). The model spectrum originates as a directional wave number spectrum, $F(K,\phi)$, with the directionally averaged form displayed here for comparison only. Model results corresponding to wave numbers $K>1.1$ (vertical, dashed line) are incorporated into the $\tilde{C}(u,v)$ function.

sequent change in coherence being well predicted by the expression for kd_H^*. The scaling relations for vertical coherence are less reliable for ASIAEX acquisition geometries owing to the lower sea-surface grazing angles involved. Nevertheless, the scaling relations provide rough approximations for d_V^*, d_H^*, and d_{HL}^* that are sufficiently close to those associated with full modeling of vertical spatial coherence to have predictive power. Rounding off the numbers for the ASIAEX acquisition geometries and surface conditions discussed here, for the 20-kHz source at depth 25 m and range 0.5 km, a coherence volume of $(2d_V^*)(2d_H^*)(2d_{HL}^*)$ exists for receivers placed between 25 and 50 m, which applies to the spatial coherence of a sound field arriving via the surface bounce channel. The volume consists of a vertical layer 0.5 m thick by 3 m in each of the two horizontal dimensions, or 4–5 m.[3] For the 8-kHz source these dimensions should be increased by a factor of 2.5, representing the ratio of the two frequencies.

ACKNOWLEDGMENT

This work is supported by the Office of Naval Research, Code 321 Ocean Acoustics Program, via Contract No. N00014-96-1-0325.

APPENDIX: TWO-DIMENSIONAL AUTOCORRELATION FUNCTION

Central to the computation of the sea-surface bistatic cross section is an estimate of the two-dimensional autocorrelation function of sea-surface waveheight variation, $C(u,v)$, as a function of Cartesian spatial lag variables u and v, where in our usage u is aligned with the acoustic line-of-site axis. We use a combination of observation and modeling to obtain $C(u,v)$. This method has evolved from one published previously[1] to utilize recent advances in the modeling of sea-surface wave directional spectra. The observable from the wave buoy is an estimate of the directional frequency spectrum expressed as a function of wave temporal frequency f and wave direction (from) ϕ. The buoy data are converted to a directional wave number spectrum that is a function of surface wave number K and ϕ; call this spectral form $F_L(K,\phi)$, where the subscript stands for "low pass." The gravity wave dispersion relation, $K=(2\pi f)^2/g$, where g is gravitational acceleration, is used in the conversion to map f to K. The maximum frequency utilized in this procedure is 0.52 Hz, rather than 0.64 Hz, as the buoy does not transmit directional data associated with higher frequency bins if the data are not above a certain noise threshold (as was the case here) and thus the maximum K in $F_L(K,\phi)$ is 1.1. Figure 7(a) shows the directional frequency spectrum measured contemporaneously with the acoustic measurements. During this period the wind and wave fields originate primarily from the north.

For the higher wave numbers not sensed by the buoy, we employ a model from Plant[18] (the D-spectrum), which gives a directional wave number spectrum as a function of fetch and wind speed. This model is computed for a wind speed of 7 m/s and for fully developed conditions in view of the O(100)-km fetch in the East China Sea, and the model result matches well with the buoy data at the necessary transition point in the K domain. This is shown in Fig. 7(b), where for this display both the buoy and model directional wave number spectra are collapsed to their directionally averaged wave number spectral forms, $F(K)$. We use the model result only where observations by the buoy are lacking. Thus, the model result is high-pass filtered by removing energy carried by wave numbers $K\leqslant 1.1$ and the model remains supported by

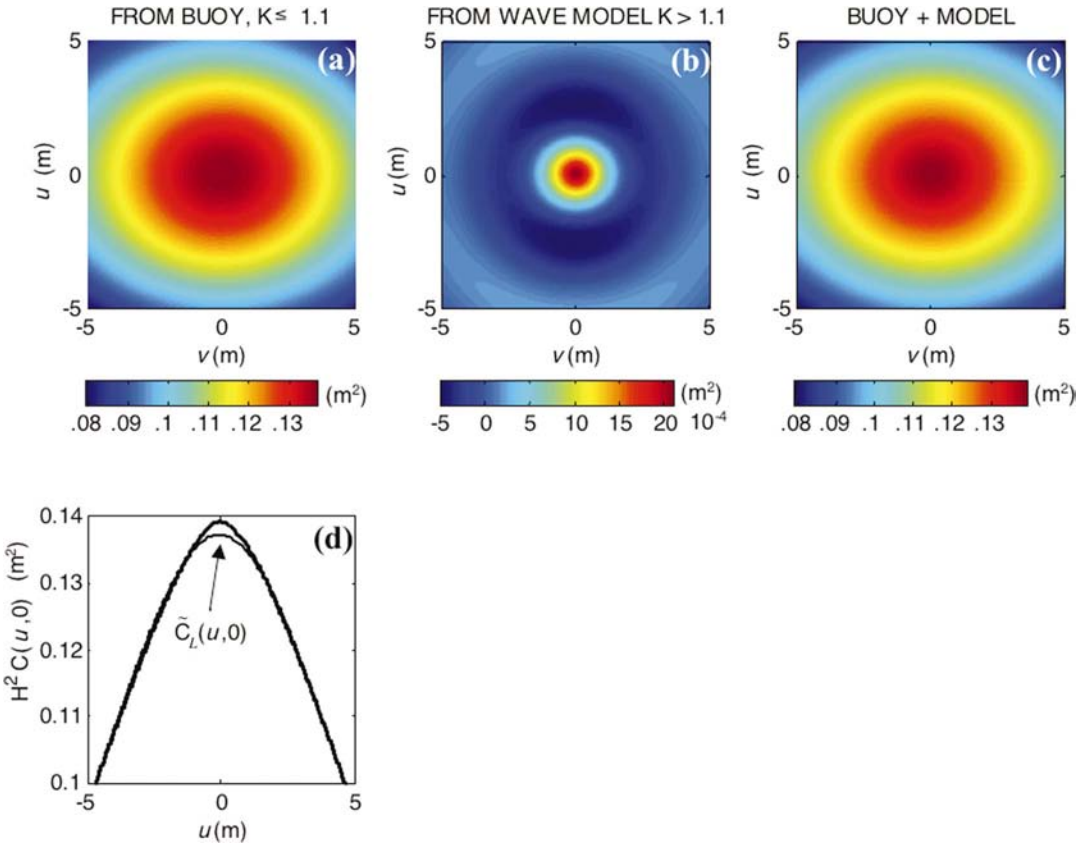


FIG. 8. Estimates of the two-dimensional autocovariance function of sea-surface waveheight variation, $\tilde{C}(u,v)$, based on a combination of buoy observations and wave model (Ref. 18). (a) Low-pass, $\tilde{C}_L(u,v)$, derived from buoy observations; (b) high-pass, $\tilde{C}_H(u,v)$, derived from model; (c) the sum $\tilde{C}(u,v)=\tilde{C}_L(u,v)+\tilde{C}_H(u,v)$; and (d) views of $\tilde{C}_L(u,v)$ (arrow) and $\tilde{C}(u,v)$ along the $v=0$ axis.

wave numbers up to 500; call this $F_H(K,\phi)$, where the subscript stands for "high pass."

Next, both low- and high-pass spectral forms are converted to their equivalent autocovariance functions via

$$\tilde{C}_{L,H}(u,v)=\int_{-\pi}^{\pi}d\phi\int_K dK F_{L,H}(K,\phi)\times\cos(K[u\cos\phi+v\sin\phi]), \tag{A1}$$

where $\tilde{C}_{L,H}$ is either the low- or high-pass autocovariance function, and integral over K corresponds to the respective ranges for each spectral form. The estimate of $C(u,v)$ used in Eq. (1) of Ref. 1 to compute the bistatic cross section is generated by taking $\tilde{C}(u,v)=\tilde{C}_L(u,v)+\tilde{C}_H(u,v)$, then normalizing $\tilde{C}(u,v)$ with the total mean-square waveheight H^2 such that $C(0,0)=1$.

Figure 8 shows the low- and high-pass portions of the autocovariance functions and their sum. During the acoustic measurements, the MELVILLE (source) was positioned nearly precisely due north of the receiving VLA system. Given the orthogonal spatial lag variables u and v are defined such that $C(u,0)$ describes roughness correlation along a line parallel to that connecting the acoustic source and receiver, then the u coordinate in this case corresponds to an up/down-wind direction and the v coordinate corresponds to a cross-wind direction. The elliptical nature of both $\tilde{C}_L(u,v)$ and $\tilde{C}(u,v)$ is consistent with expected lower correlation along up/down-wind direction, and the wave slope variance also takes on an elliptical variation but rotated 90°. With $\tilde{C}_H(u,v)$ the elliptical property is less evident owing to its support consisting of higher wave numbers that are less directive. The faint patterns at the top and bottom of $\tilde{C}_H(u,v)$ are zero crossings associated with the minimum supporting wave numbers close to 1.1 radians/m. [For $\tilde{C}_L(u,v)$ zero crossings associated with the dominant wavelength of O(100) m are not shown in the scale of the plot.] Plots of $\tilde{C}_L(u,0)$ and $\tilde{C}(u,0)$ along the up/down-wind direction [Fig. 8(d)] show that the primary effect of adding $\tilde{C}_H(u,v)$ to $\tilde{C}_L(u,v)$ is to increase the curvature near the origin with little change in total waveheight variance. The mean-square slope along this direction S_u^2, equal to $[\partial^2\tilde{C}(u,0)]/\partial u^2$ evaluated at $u=0$, is 0.0157. The similarly computed mean-square slope in the cross-wind direction S_v^2 is 0.0132, giving a total mean-square slope of 0.0289 with ratio S_v^2/S_u^2 equal to 0.84.

Two remarks are in order concerning the $\tilde{C}(u,v)$ function and its variance-normalized counterpart $C(u,v)$. First, there is little difference between the model results for spatial coherence in Fig. 5 and ones based on a directionally averaged, or isotropic, version of $\tilde{C}(u,v)$. This is anticipated from previous work.[1] Model curves for $|\Gamma_V|$ based on a directionally averaged $\tilde{C}(u,v)$ will, in each case, be slightly above those shown in Fig. 5, with two such curves slightly overlapping on the scale of the figure. In short, we cannot

resolve effects due to wave directivity in these estimates of spatial coherence, each of which is based on 60 independent ping receptions.

Second, although the $\tilde{C}(u,v)$ function is supported by surface wave numbers K up to 500 radian/m, they need not be for the purposes of computing spatial coherence. As discussed in Ref. 2, length scale support by wave numbers up to about $k/4$ is essentially all that is needed to reproduce the physics of forward scattering governing spatial coherence in the order 10–100-kHz frequency range. At 20 kHz $k/4$ is about 21 radians/m, and redoing the spatial coherence calculations using a low-pass version $\tilde{C}(u,v)$ based on surface wave numbers $K \leqslant 21$, yields essentially the same results for spatial coherence. In Ref. 2, such a mean-square slope computed for $K \leqslant k/4$ was called the effective mean-square slope of the large-scale waves, or S_L^2. A simple formula is provided in Ref. 2 that predicts $S_L = 0.146$ for a wind speed of 7 m/s and frequency of 20 kHz, which is quite consistent with the estimate of $S_L = 0.152$ based on the low-pass version $\tilde{C}(u,v)$ derived here.

Finally, what the $\tilde{C}(u,v)$ function based on full wave number support and wave directivity does provide is the capability to predict bistatic sea-surface scattering for scattering regimes ranging from the near-specular regime to the Bragg scattering regime, and to evaluate the effects of wave directivity for a given acoustic frequency, using the small slope approximation.[1] This is a separate subject beyond the scope of this work.

[1] P. H. Dahl, "On bistatic sea surface scattering: Field measurements and modeling," J. Acoust. Soc. Am. **105**, 2155–2169 (1999).

[2] P. H. Dahl, "High-frequency forward scattering from the sea surface: The characteristic scales of time and angle spreading," IEEE J. Ocean. Eng. **26**, 141–151 (2001).

[3] W. M. Carey and W. B. Mosely, "Space-time processing, environmental-acoustic effects," IEEE J. Ocean. Eng. **16**, 285–301 (1991).

[4] S. M. Flatte, R. Dashen, W. H. Munk, K. M. Watson, and F. Zachariasen, *Sound Transmission Through a Fluctuating Ocean*, 2nd ed. (Cambridge University Press, New York, 1979).

[5] M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Cambridge University Press, New York, 1999), pp. 572–576.

[6] J. Pernia, *Coherence of Light*, 2nd ed. (Reidel, Boston, 1985), pp. 33–36.

[7] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy* (Wiley, New York, 1986), pp. 73–77.

[8] R. Mallart and M. Fink, "The van Cittert–Zernike theorem in pulse echo measurements," J. Acoust. Soc. Am. **90**, 2718–2727 (1991).

[9] D. R. Jackson and K. Y. Moravan, "Horizontal spatial coherence in ocean reverberation," J. Acoust. Soc. Am. **75**, 428–436 (1984).

[10] V. Hesany, W. J. Plant, and W. C. Keller, "The normalized radar cross section of the sea at 10° Incidence," IEEE Trans. Geosci. Remote Sens. **38**, 64–72 (2000).

[11] J. R. Apel, "An improved model of the ocean surface wave vector spectrum and its effects on radar backscatter," J. Geophys. Res. **99**, 8 16269–16291 (1994).

[12] D. E. Barrick, "Rough surface scattering based on the specular point theory," IEEE Trans. Antennas Propag. **AP-16**, 449–454 (1968).

[13] P. H. Dahl, "On the spatial coherence and angular spreading of sound forward scattered from the sea surface: Measurements and interpretive model," J. Acoust. Soc. Am. **100**, 748–758 (1996).

[14] S. T. McDaniel, "Spatial covariance and adaptive beam forming of high-frequency acoustic signals forward scattered from the sea surface," IEEE J. Ocean. Eng. **16**, 415–419 (1991).

[15] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, 2nd ed. (Griffen, London, 1963), Vol. 1, p. 236.

[16] B. Efron and T. Tibshirani, "Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy," Stat. Sci. **1**, 54–77 (1986).

[17] P. H. Dahl, "The contribution of bubbles to high-frequency sea surface backscatter: A 24-h time series of field measurements," J. Acoust. Soc. Am. **113**, 769–780 (2003).

[18] W. J. Plant, "A stochastic, multiscale model of microwave backscatter from the ocean," J. Geophys. Res. **107**, 10.129/2001 JC000909 (2002).

Multiple focusing with adaptive time-reversal mirror

J. S. Kim[a] and K. C. Shin
Division of Ocean Engineering, Korea Maritime University, Busan, 606-791, Korea

(Received 31 October 2002; revised 14 November 2003; accepted 14 November 2003)

Recently, adaptivity was introduced to time-reversal mirror to steer the nulls, and referred to as an adaptive time-reversal mirror (ATRM) [J. S. Kim, H. C. Song, and W. A. Kuperman, J. Acoust. Soc. Am. **109**, 1817–1825 (2001)]. In this study, ATRM is extended to simultaneous multiple focusing in an ocean waveguide. The multiple focusing is achieved by imposing a set of constraints in the formulation to find the weight vectors. The algorithm is applied to the long-range underwater acoustic communication to show, via simulation, that the simultaneous pulse compression at multiple receiving locations is achieved. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1640389]

PACS numbers: 43.30.Vh, 43.20.Fn [RAS] Pages: 600–606

I. INTRODUCTION

It is well known that a time-reversal mirror (TRM) uses the received signal from a probe source to refocus the signal at the probe source location by back-propagating the time-reversed version of the received signal. Since TRM has been demonstrated in ultrasonics[1,2] and in ocean acoustics,[3–5] a specific application of TRM to achieve low error bit rate in underwater communication was successfully demonstrated.[6]

Recently, the concept of an adaptive weighting on the TRM array before back-propagation was introduced to time-reversal mirror and referred to as an adaptive time-reversal mirror (ATRM).[7] Using optimization theory with constraints,[8–10] an expression for a null steering vector can be formulated at the arbitrary location, assuming that the cross-spectral density matrix (CSDM) for the null location is known. Also the applicability of ATRM has been demonstrated in selective focusing onto a weak scatterer among the two scatterers.[7]

In the present study, ATRM is extended to simultaneous multiple focusing in an ocean waveguide by imposing a set of constraints in the formulation of the weight vectors. Via simulation the developed algorithm is applied to the pulse compression process at the multiple receiving locations in the long-range underwater acoustic communication, where the characteristics of the channel is reverberating and dispersive resulting in a severe signal distortion.

In Sec. II, the formulation of TRM and ATRM is reviewed, and ATRM is extended to the simultaneous multiple focusing algorithm. Simulation results are presented in Sec. III for TRM and ATRM in an ocean waveguide in the context of underwater communication. A summary of the work presented is found in Sec. IV.

II. THEORY

A. Review of TRM

In order to review the adaptive TRM, we follow the convention in Kim.[7] As described in Fig. 1, the phase-conjugate field at the field location $\vec{r}$ in frequency domain is written as

$$p(\vec{r}) = \sum_{i=1}^{N} g^*(\vec{r}_i|\vec{r}_{\mathrm{ps}})\, g(\vec{r}|\vec{r}_i) = \mathbf{g}^+(\mathbf{r}_{\mathrm{array}}|\vec{r}_{\mathrm{ps}})\, \mathbf{g}(\vec{r}|\mathbf{r}_{\mathrm{array}}), \tag{1}$$

where $g(\vec{r}_i|\vec{r}_{\mathrm{ps}})$ represents the received acoustic pressure at the ith array element location $\vec{r}_i$ propagated from the probe source position $\vec{r}_{\mathrm{ps}}$. Likewise, $g(\vec{r}|\vec{r}_i)$ represents the field propagated from the ith array element location $\vec{r}_i$ to the arbitrary receiver location $\vec{r}$ as shown in Fig. 1, where N is the number of array elements and superscripts $(\ )^*$ and $(\ )^+$ denote complex conjugate and Hermitian transpose, respectively. In a vector notation, $\mathbf{g}$ and $\mathbf{r}_{\mathrm{array}}$ are $(N\times 1)$ column vectors. Note that the position vectors are written in *italic* letters with arrows and the column vectors and matrices are written in boldface letters.

B. Review of ATRM

Introducing a signal vector for back-propagation

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{N-1} \\ w_N \end{bmatrix}, \tag{2}$$

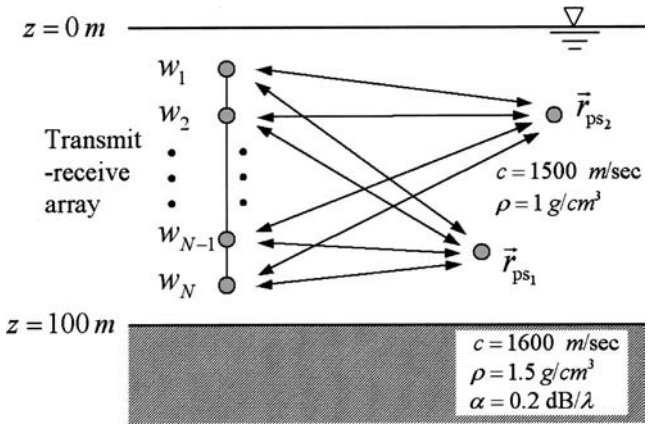


FIG. 1. Description of Pekeris waveguide for TRM.

[a]Author to whom correspondence should be addressed. Electronic mail: jskim@hhu.ac.kr

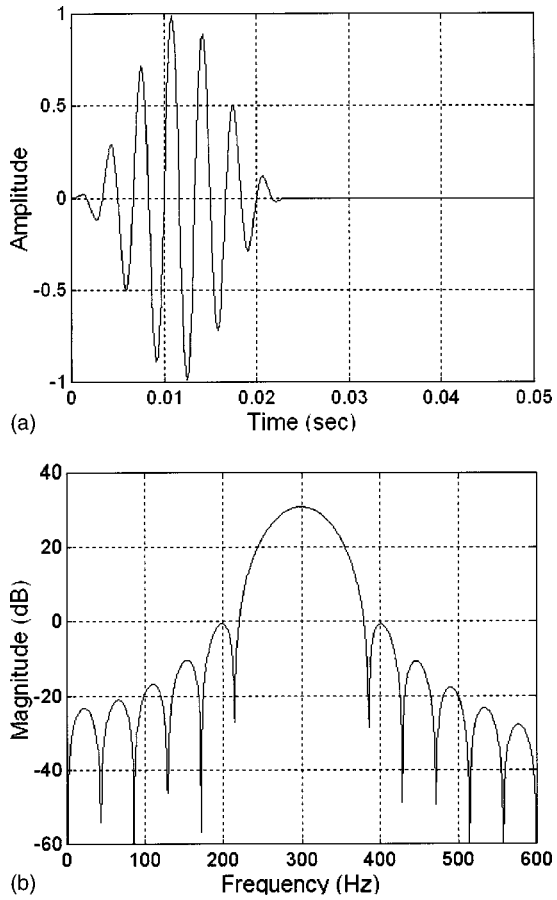


FIG. 2. Probe source (a) signal and (b) spectrum.

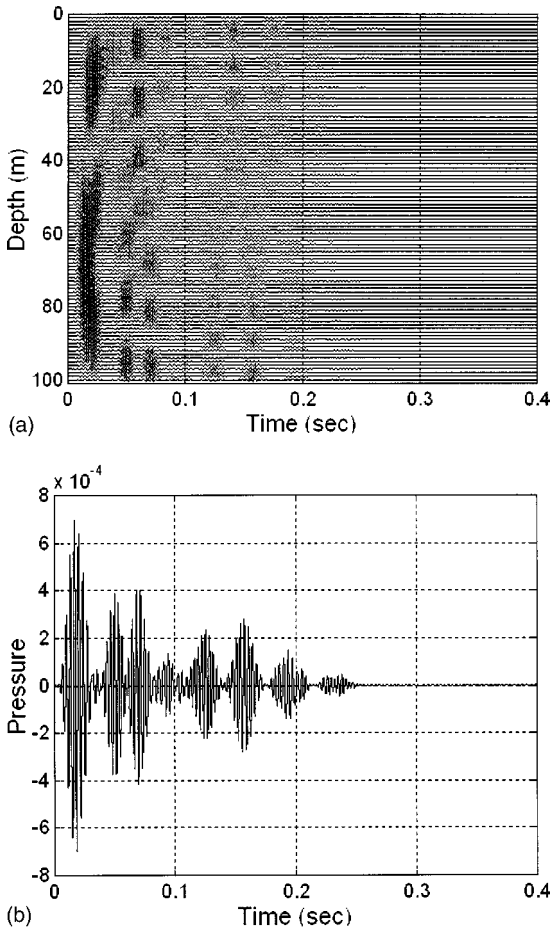


FIG. 3. The time series received at the array with a probe source at *A* (5800 m, 80 m).

Eq. (1) can, more generally, be written as

$$p(\vec{r}) = \sum_{i=1}^{N} w_i^* g(\vec{r}|\vec{r}_i) = \mathbf{w}^+ \mathbf{g}(\vec{r}|\mathbf{r}_{\text{array}}), \tag{3}$$

where **w** reduces to **g** in a conventional time-reversal mirror.

In the previous study,[7] the formulation for the signal vector **w** to steer the nulls was derived with an application to the selective focusing to a weak target. The formulation is based on the optimization problem of the following objective function:

$$\min_{\mathbf{w}} \mathbf{w}^+ \mathbf{K} \mathbf{w}, \tag{4}$$

where the cross-spectral density matrix **K** is expressed as

$$\mathbf{K} = \mathbf{g}(\vec{r}_{\text{ps}}|\mathbf{r}_{\text{array}})\mathbf{g}^+(\vec{r}_{\text{ps}}|\mathbf{r}_{\text{array}}) + \mathbf{g}(\vec{r}_{\text{n}}|\mathbf{r}_{\text{array}})\mathbf{g}^+(\vec{r}_{\text{n}}|\mathbf{r}_{\text{array}}) \tag{5}$$

with the distortionless response constraint at the focal location, which can be expressed as

$$\mathbf{w}^+ \mathbf{g}(\vec{r}_{\text{ps}}|\mathbf{r}_{\text{array}}) = 1. \tag{6}$$

The nulling location is denoted as $\vec{r}_{\text{n}}$. This optimization problem has a well-known solution referred to as the minimum variance in the adaptive array signal processing.[8–10]

C. Multiple focusing with ATRM

In the present study, a formulation for the signal vector **w** to have the multiple focusing points is derived and applied to the long-range underwater communication in the ocean waveguide in order to demonstrate that the pulse compression at multiple receiving locations is realized in a highly dispersive multipath environment.

Weight vector **w** in Eq. (3) gives a distortionless response for one probe source only. We want to achieve simultaneous multiple focusing with distortionless response in multiple probe source locations. For this purpose, the multiple linear constraint ATRM is applied. In order to be more explicit in the formulation, the multiple linear constraint can be expressed as

$$\mathbf{M}^+ \mathbf{w} = \mathbf{c}, \tag{7}$$

where **M** is the $(N \times P)$ constraint matrix which is consisted of signal vectors at probe source locations, **w** is the $(N \times 1)$ weight vector, and **c** is the $(P \times 1)$ response vector. It is noted that N is the number of array elements, and P is the number of locations to be focused (or, self-equalized in communication) when the weight vector **w** is back-propagated.

We formally state the general optimization problem of the following objective function:

$$\min_{\mathbf{w}} \mathbf{w}^+ \mathbf{K} \mathbf{w} \quad \text{subject to} \quad \mathbf{M}^+ \mathbf{w} = \mathbf{c}. \tag{8}$$

The solution to this problem is also well known. Since the derivation is simple, it is briefly explained. Forming the functional with the multiple linear equality constraints gives[8]

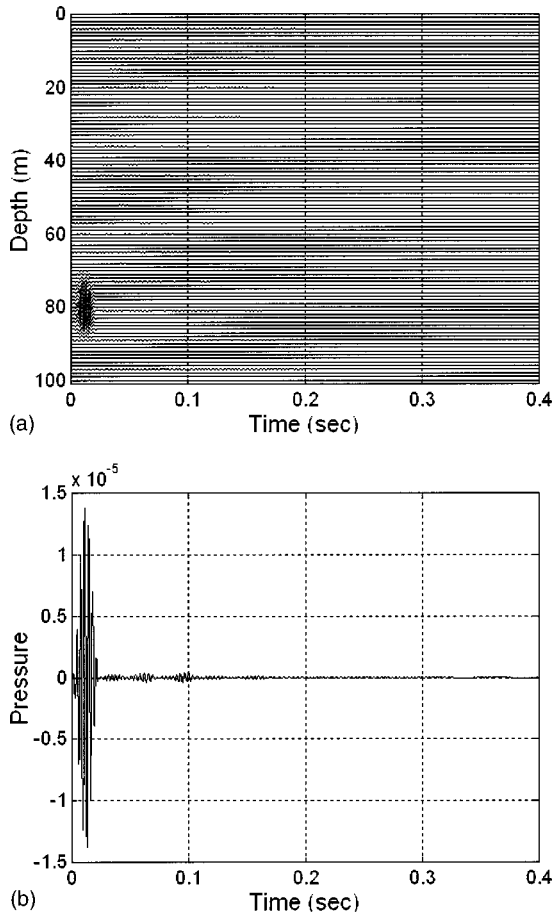


FIG. 4. Simulated time series focused on the probe source location at *A* (5800 m, 80 m), (a) depth-stacked time series and (b) time series at 80 m depth.

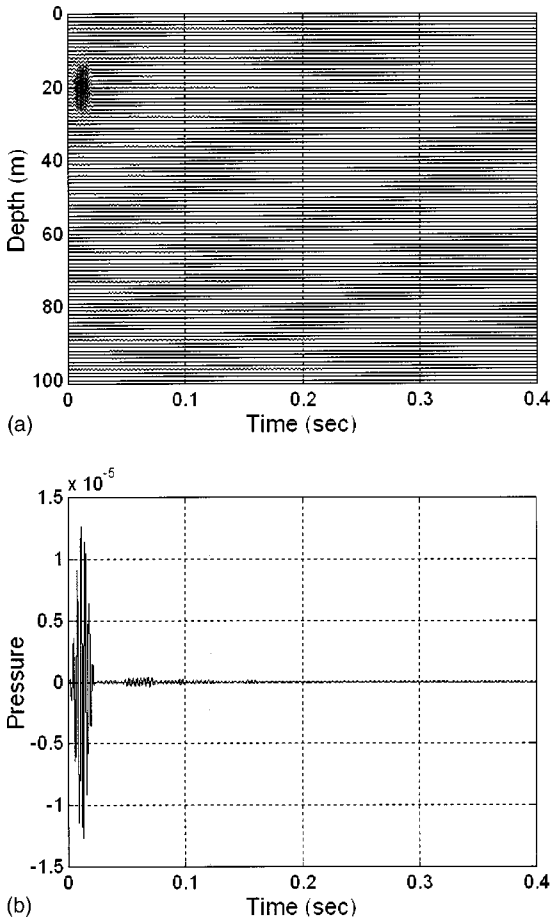


FIG. 5. Simulated time series focused on the probe source location at *B* (6200 m, 20 m), (a) depth-stacked time series and (b) time series at 20 m depth.

$$\mathbf{F}=\mathbf{w}^{+}\mathbf{K}\mathbf{w}-\lambda(\mathbf{M}^{H}\mathbf{w}-\mathbf{c}). \tag{9}$$

Taking the derivative with respect to the weight vector, and solving for λ, gives the solution for the weight vector

$$\mathbf{w}=\mathbf{K}^{-1}\mathbf{M}[\mathbf{M}^{+}\mathbf{K}^{-1}\mathbf{M}]^{-1}\mathbf{c}. \tag{10}$$

The response vector associated with constraint matrix has 1 for its all element, corresponding to the probe sources location. However, the value of the response matrix can be prescribed to monitor the relative strength of the focused field.

For the numerical experiment in Sec. III with Pekeris waveguide shown in Fig. 1, the CSDM is defined as

$$\mathbf{K}=\mathbf{g}(\vec{r}_{\mathrm{ps1}}|\mathbf{r}_{\mathrm{array}})\mathbf{g}^{+}(\vec{r}_{\mathrm{ps1}}|\mathbf{r}_{\mathrm{array}}) + \mathbf{g}(\vec{r}_{\mathrm{ps2}}|\mathbf{r}_{\mathrm{array}})\mathbf{g}^{+}(\vec{r}_{\mathrm{ps2}}|\mathbf{r}_{\mathrm{array}}) \tag{11}$$

and the constraints for distortionless response are

$$\mathbf{w}^{+}\mathbf{g}(\vec{r}_{\mathrm{ps1}}|\mathbf{r}_{\mathrm{array}})=1 \quad \text{and} \quad \mathbf{w}^{+}\mathbf{g}(\vec{r}_{\mathrm{ps2}}|\mathbf{r}_{\mathrm{array}})=1. \tag{12}$$

In matrix form, the multiple constraints in Eq. (12) is expressed as $\mathbf{M}^{+}\mathbf{w}=\mathbf{c}$, where $\mathbf{M}=[\mathbf{g}(\vec{r}_{\mathrm{ps1}}|\mathbf{r}_{\mathrm{array}})\,\mathbf{g}(\vec{r}_{\mathrm{ps2}}|\mathbf{r}_{\mathrm{array}})]$ and $\mathbf{c}=[\begin{smallmatrix}1\\1\end{smallmatrix}]$. Finally, the signal vector $\mathbf{w}$ for back-propagation is found from Eq. (10).

III. MULTIPLE FOCUSING IN AN OCEAN WAVEGUIDE

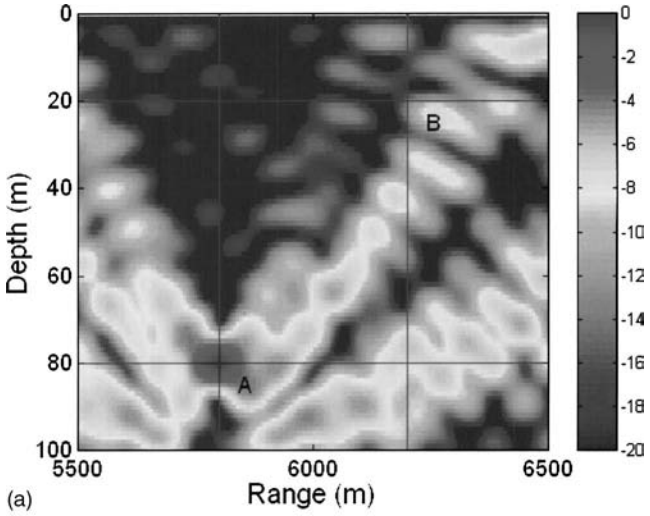
In this section, the multiple focusing is demonstrated in Pekeris waveguide which is shown in Fig. 1. A time-reversal mirror is used to focus on the probe sources location in Sec. III A. In Sec. III B, A superimposed phase-conjugate field by two probe sources and its time series are shown. In Sec. III C, it is demonstrated that focusing in multiple probe source locations have a distortionless response.

As seen in Fig. 2 the transmitted probe pulse has a center frequency $f_c=300$ Hz and ping duration $\tau=23.33$ ms. The sampling frequency is 3600 Hz, FFT size is 8192, so that the time duration is 2.3 s long. This time duration is long enough to prevent aliasing in time as a pulse propagates in the ocean and stretches in time due to multipath propagation and dispersion.

The probe sources are located at *A*, $\vec{r}_{\mathrm{ps}_1}=(5800\,\mathrm{m},\ 80\,\mathrm{m})$ and *B*, $\vec{r}_{\mathrm{ps}_2}=(6200\,\mathrm{m},\ 20\,\mathrm{m})$, respectively. These focal locations are separated to satisfy the diffraction limit. The vertical transmit/receive array consists of 17 elements with an interelement spacing of 5 m ranging from 10 m to 90 m in depth. The description of an ocean waveguide is shown in Fig. 1. As a propagation model, KRAKEN[11] was used throughout the simulation.

A. Time-reversal mirror at each probe source locations

For the purpose of developing the multiple focusing process, the phase-conjugate field with one probe source is examined. Figure 3(a) shows the depth-stacked time series re-



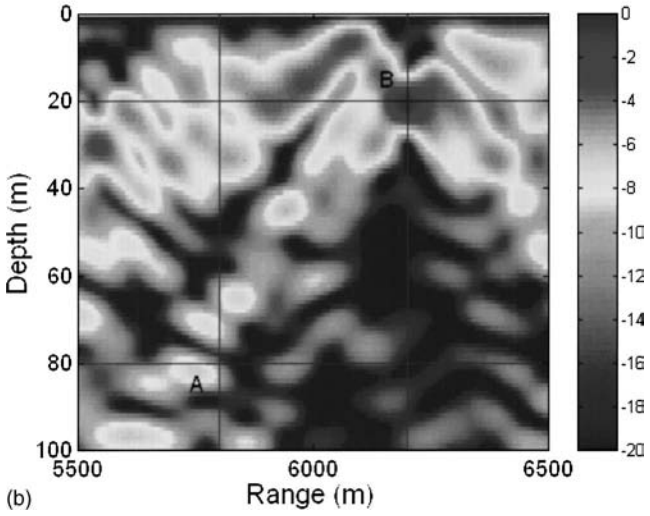


FIG. 6. Single frequency phase conjugation experiment at 300 Hz, (a) focusing at A (5800 m, 80 m) and (b) focusing at B (6200 m, 20 m).

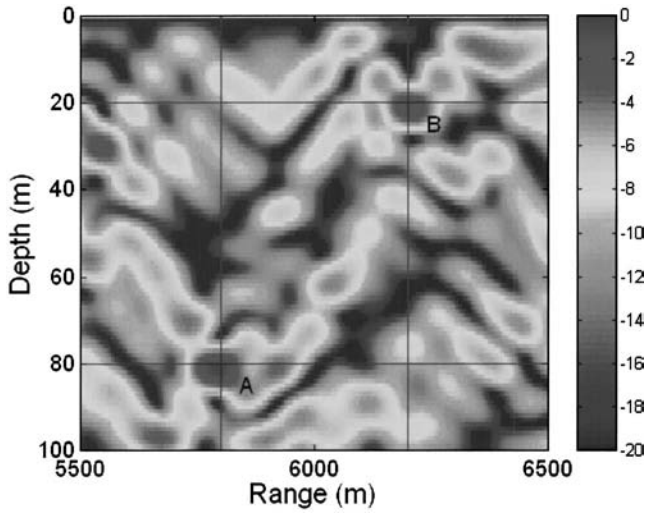


FIG. 7. Focusing at point A and B with phase conjugation based on superposition.

ceived at the range of 5800 m, and the signal received at the array elements at the depth of 80 m from the source location A is shown in Fig. 3(b). As expected, the signal is spread in time due to multipath and dispersion so that the signal is hardly correlated with the source signal.

When the received signal at the array is time-reversed and back-propagated, the signal focuses back to the original source location as shown in Figs. 4 and 5. Figures 4(a) and 5(a), respectively, illustrate single pulse propagation from a source located at ranges 5800 m and 6200 m away as a function of depth. The time is referenced at the start of the focused signal, since the field response is one at the focal point. The focusing is noted at 80 m and 20 m depth, of which time series are shown in Figs. 4(b) and 5(b).

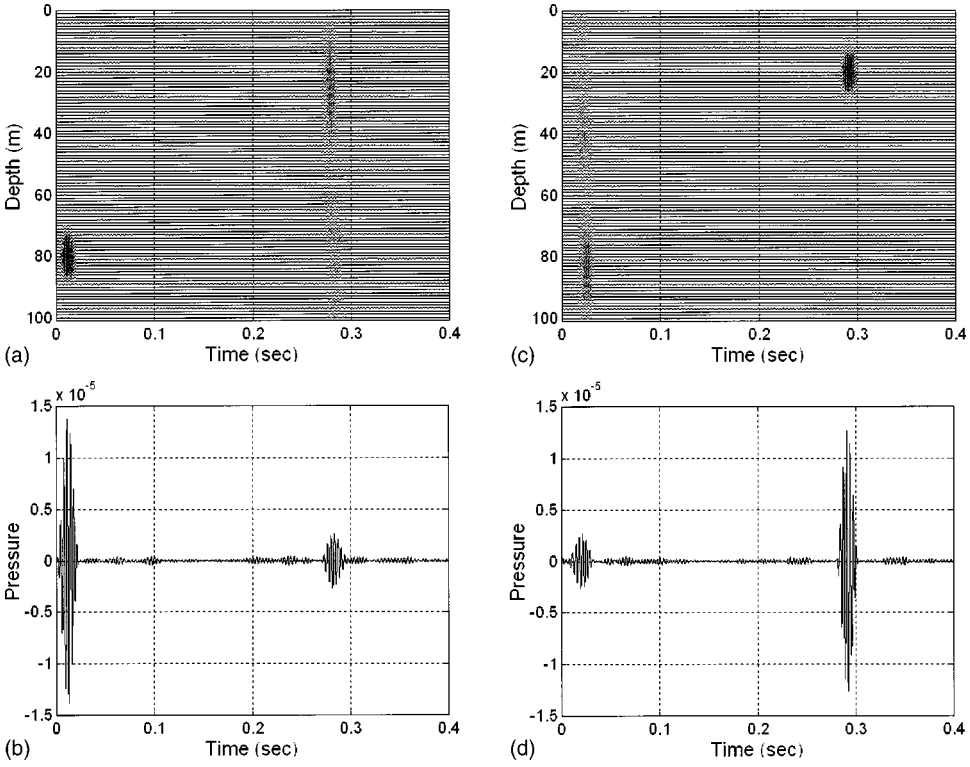


FIG. 8. Simulated back-propagated time series with two superimposed focusing: (a) depth-stacked time series at range 5800 m, (b) time series at A (5800 m, 80 m), (c) depth-stacked time series at range 6200 m, and (d) time series at B (6200 m, 20 m).

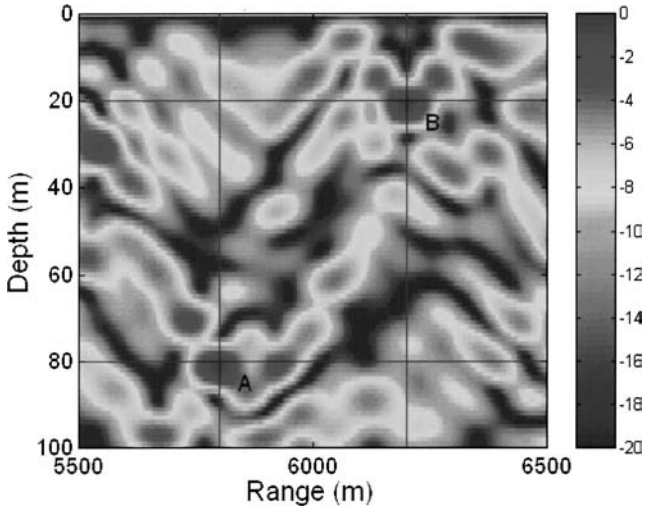


FIG. 9. Mutiple focusing with distortionless responses at points *A* and *B*.

Figure 6 shows the focused field at center frequency 300 Hz. The spatial focusing at the probe source location is clearly observable. The position *A*, $\vec{r}_{ps_1} = (5800\,\text{m},\ 80\,\text{m})$ and B, $\vec{r}_{ps_2} = (6200\,\text{m},\ 20\,\text{m})$ are focal locations in Fig. 6.

B. Superimposed TRM at two probe sources

Figure 7 shows the field at the center frequency 300 Hz, when the two weight vectors found in Sec. III A for each focal locations are superimposed and back-propagated. The depth-stacked time series are shown in Figs. 8(a) and (c) at ranges 5800 m and 6200 m, respectively. Figures 8(b) and (d) show the time series at the points *A* and *B*, when the superimposed weight vectors are back-propagated. Because of the interference between the weight vectors, the pulse compression at *A* and *B* are not fully achieved as expected so that the intersymbol interference (ISI) becomes significant. In the following Sec. III C, simultaneous focusing at *A* and *B* is demonstrated without interference.

C. TRM with distortionless response at two probe sources

When the weight vector found from Eq. (10) is back-propagated, the multiple focusing is achieved as shown in Fig. 9. The depth-stacked time series at the range of the focal locations *A* and *B* are shown in Figs. 10(a) and (c). The time series observed at the focal locations *A* and *B* are shown in Figs. 10(b) and (d). The difference between Fig. 10 and Fig. 8 is that there is no interference when ATRM is used.

In order to obtain the time series in Fig. 10, the weight vectors are calculated at each frequency bin so that the probe signal does not have to be narrow banded. For the wide band signal, the weight vectors at each frequency bin are then combined with the probe source signal specturm, and back-propagated. Finally, the received signal components at each frequency bin are transformed into the time domain by inverse fast Fourier transform (FFT) to find the time domain solution. When the probe signal is narrow banded, the weight vector can be calculated only once and used over the bandwidth. By calculating the weight vector only once, the computational time can be saved, but the temporal sidelobe level gets higher due to the mismatch off the center frequency.

As for the spatial sidelobe levels, since the multiple constraints for distortionless responses at multiple focal locations consume the degree of freedom of the weight vectors on transmit/receive array elements, the spatial sidelobe levels get higher. Although it is not easy to quantify the degradation of the peak-to-sidelobe levels, a general rule is that, when the number of foci increases, the spatial sidelobe levels get

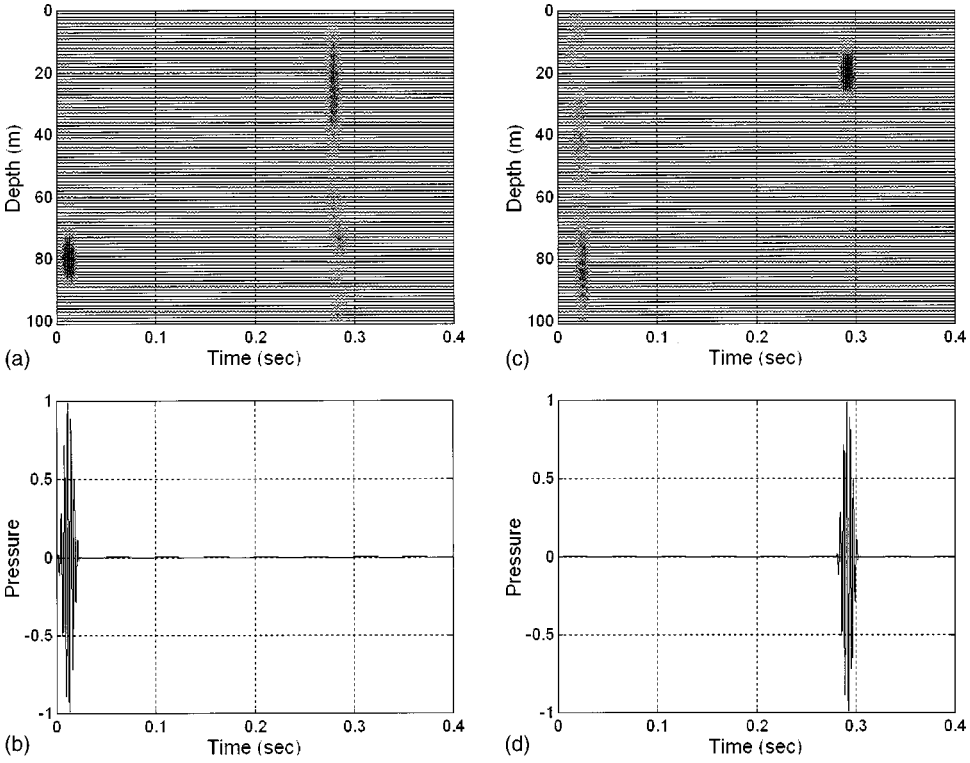


FIG. 10. Time series simultaneously focused at the two locations with ATRM, (a) depth-stacked time series at range 5800 m, (b) time series at range 5800 m and depth 80 m, (c) depth-stacked time series at range 6200 m and (d) time series at range 6200 m and depth 20 m.

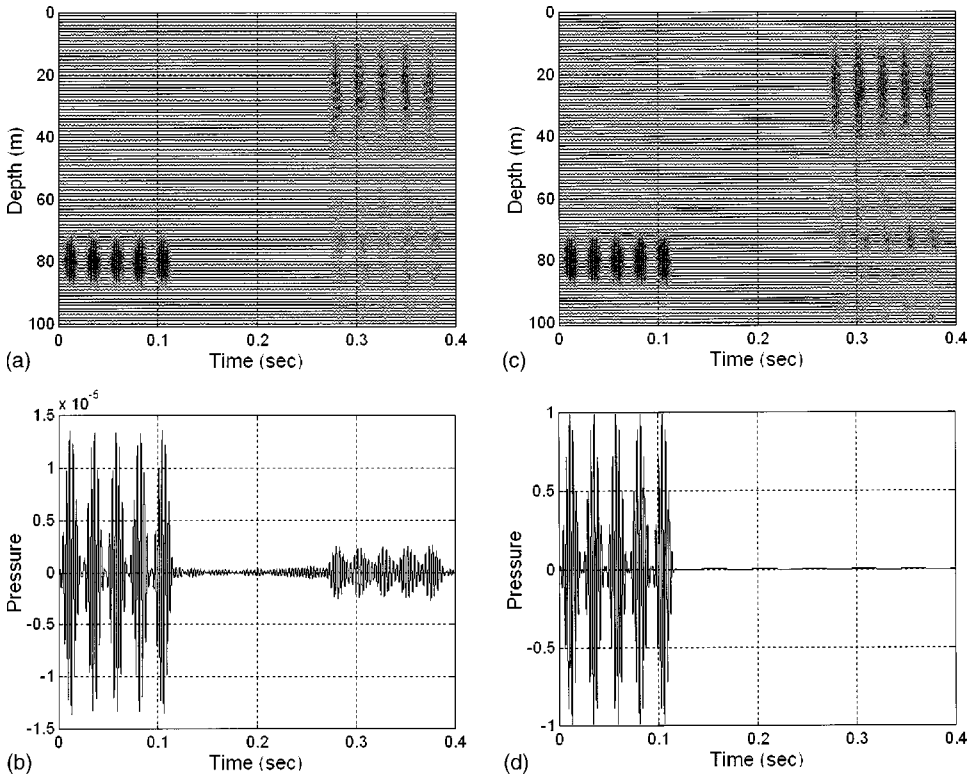


FIG. 11. BPSK pulse propagation, (a) depth-stacked time series at the range 5800 m and (b) time series at A (5800 m, 80 m). There is an intersymbol interference with superposed weight vectors, (c) depth-stacked time series at the range 5800 m with weight vector based on ATRM. There is no ISI. (d) Time series at A (5800 m, 80 m) with ATRM.

higher. Of course, the loss of degrees of freedom can be compensated by increasing the number of transmit/receive array elements.

D. Application to long-range underwater communication

The algorithm presented can be applied to the long range underwater communication in shallow water, where the signal distortion is severe and pulse compression at multiple receiving locations is desirable.

In the numerical simulation, binary phase shift keying (BPSK) is used to represent either a 1 or 0 by changing the phase of the pulse: The communication sequence length is 5 bits lasting 0.116 s. Figures 11(b) and (d) show that there exists crosstalk and intersymbol interference when the superposed weight vectors are individually calculated, while the multiple focusing at A and B with ATRM in Figs. 11(a) and (c) show no ISI and crosstalk.

Figure 12 shows the constellation plots for 150 bit sequences. When TRM alone is used, the performance is degraded due to ISI and crosstalk as shown in Fig. 12(a). However, when ATRM is used, the performance virtually converges into a point as shown in Fig. 12(b). The effect of noise is excluded in Fig. 12 in order to emphasize the applicability of multiple focusing with ATRM. The degree of constellation in Fig. 12(a) depends on the sidelobe level of the point B while focusing on the point A in Fig. 6(b), which represents the level of crosstalk between the two receiving locations, A and B.

When inverting the cross-spectral density matrix $\mathbf{K}$ in Eq. (10), the rank deficiency and the effect of noise can make the inversion unstable. The common remedy for this problem is diagonal loading. A systematic way for diagonal loading is

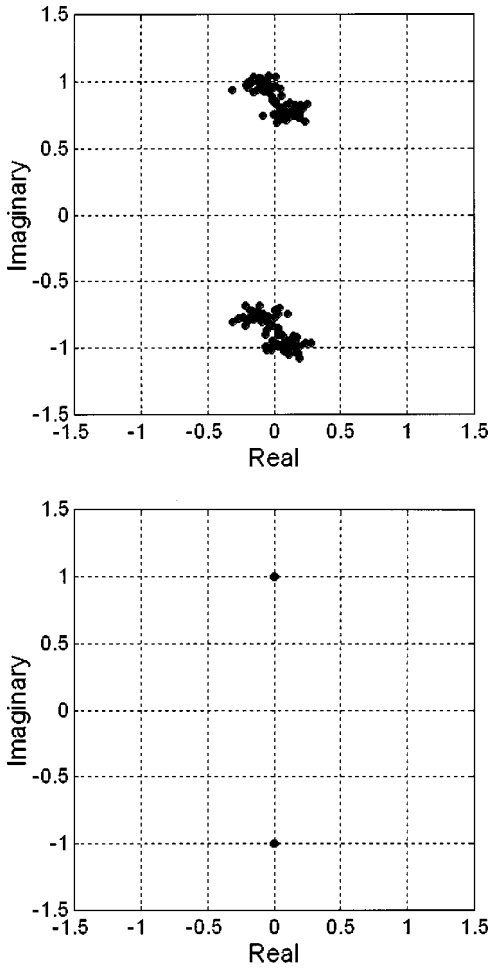


FIG. 12. Constellation plots, (a) with ISI and crosstalks based on TRM and (b) based on ATRM.

white noise gain (WNG).[12] In this paper, the white noise constraint (WNC) method[12] is used to make the inversion robust to the noise.

IV. CONCLUSION

Adaptive time-reversal mirror has been applied to the pulse compression at the multiple receiving locations in a highly dispersive and reverberating channel. The theory and formulation to find the weight vector to be back-propagated is presented, and simulation results are shown for the long range underwater communication in shallow water. It is shown that the multiple focusing (or, pulse compression at the multiple receiving locations in communication) can be achieved.

ACKNOWLEDGMENT

This work was supported by Grant No. R05-2002-000-01097-0 from the Basic Research Program of the Korea Science and Engineering Foundation.

[1] M. Fink, ''Time-reversal mirror,'' J. Phys. D **26**, 1330–1350 (1993).

[2] M. Fink, ''Time reversed acoustics,'' Phys. Today **50**, 34–40 (1997).

[3] D. R. Jackson and D. R. Dowling, ''Phase conjugation in underwater acoustics,'' J. Acoust. Soc. Am. **89**, 171–181 (1991).

[4] W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, ''Phase conjugation in the ocean: Experimental demonstration of an acoustic time reversal mirror,'' J. Acoust. Soc. Am. **103**, 25–40 (1998).

[5] W. S. Hodgkiss, H. C. Song, W. A. Kuperman, T. Akal, C. Ferla, and D. R. Jackson, ''A long-range and variable focus phase-conjugation experiment in shallow water,'' J. Acoust. Soc. Am. **103**, 25–40 (1999).

[6] G. F. Edelmann, W. S. Hodgkiss, S. Kim, W. A. Kuperman, H. C. Song, and T. Akal, ''Underwater acoustic communication using time-reversal,'' Proceeding of Oceans 2001, pp. 793–798.

[7] J. S. Kim, H. C. Song, and W. A. Kuperman, ''Adaptive time-reversal mirror,'' J. Acoust. Soc. Am. **109**, 1817–1825 (2001).

[8] D. H. Johnson and D. E. Dudgeon, in *Array Signal Processing-Concepts and Techniques*, Prentice-Hall Signal Processing Series, edited by A. V. Oppenheim (Prentice-Hall, Englewood Cliffs, NJ, 1993).

[9] H. Cox, ''Robust adaptive beamforming,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-35**, 1365–1376 (1987).

[10] J. Capon, ''High resolution frequency wavenumber spectrum analysis,'' Proc. IEEE **58**, 1408–1418 (1969).

[11] M. B. Porter, ''The KRAKEN Normal Mode Program,'' SACLANT Undersea Research Center SM-245, La Spezia, Italy, 1991.

[12] R. A. Gramann, ABF algorithms implemented at ARL:UT, ARL-TL-EV-92-31, May 1992.

Adjoint modeling for acoustic inversion[a)]

Paul Hursky[b)] and Michael B. Porter
Science Applications International Corporation, 10260 Campus Point Drive, San Diego, California 92121

B. D. Cornuelle, W. S. Hodgkiss, and W. A. Kuperman
Scripps Institution of Oceanography, University of California in San Diego, 9500 Gilman Drive, La Jolla, California 92093-0238

(Received 12 September 2002; accepted for publication 4 November 2003)

The use of adjoint modeling for acoustic inversion is investigated. An adjoint model is derived from a linearized forward propagation model to propagate data-model misfit at the observation points back through the medium to the medium perturbations not being accounted for in the model. This adjoint model can be used to aid in inverting for these unaccounted medium perturbations. Adjoint methods are being applied to a variety of inversion problems, but have not drawn much attention from the underwater acoustic community. This paper presents an application of adjoint methods to acoustic inversion. Inversions are demonstrated in simulation for both range-independent and range-dependent sound speed profiles using the adjoint of a parabolic equation model. Sensitivity and error analyses are discussed showing how the adjoint model enables calculations to be performed in the space of observations, rather than the often much larger space of model parameters. Using an adjoint model enables directions of steepest descent in the model parameters (what we invert for) to be calculated using far fewer modeling runs than if a forward model only were used. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1636760]

PACS numbers: 43.30.Pe, 43.60.Pt, 43.60.Rw [WMS] Pages: 607–619

I. INTRODUCTION

This paper investigates the use of adjoint modeling for acoustic inversion. Typically, in inversion problems a forward model driven by presumed medium properties is used to propagate a field from the source to the observation points, where the field predicted by the forward model is compared with the measured field there. The actual medium is not known and is typically a perturbation of the presumed medium, to be calculated by an inversion process so that subsequent forward modeling can be improved. If the properties of the medium used to drive the forward model match the actual medium properties, then the data-model misfit at the observation points is small. If not, these presumed medium properties are adjusted and the forward modeling repeated until the data-model misfit has been reduced. Note that the emphasis in this process is on the mapping from the unknown medium properties to the observations, when it is the medium properties that we want to estimate. What we really want is some way to propagate the data-model misfit from our observation points directly back to the points in the medium where we need to make adjustments. This is what an adjoint model does. It back-propagates an adjoint field, initialized to the data-model misfit at the observation points, back to the points in the medium where a correction to the presumed medium properties can be made to reduce the data-model misfit.

An adjoint model has particular relevance to inversion problems in which the unknown parameter space is much greater than the observation space. Methods based on adjoint models have been used in geophysical inversion,[1] electromagnetic tomography[2] and oceanography.[3–5] An adjoint approach to acoustic inverse scattering problems has also been explored, on a theoretical level, with emphasis on variational methods.[6] Other recent work in acoustics includes using adjoint methods to invert for bottom boundary conditions in the ocean waveguide[7] and using an adjoint approach with normal modes to compute derivatives with respect to environmental parameters.[8] In this paper we show how the technique can be used to invert for medium properties in underwater acoustic problems, demonstrating an adjoint-based inversion for range-independent and range-dependent sound speed profiles.

To understand the benefits of the approach for acoustic inversion, it is useful to review the commonly used alternatives. A standard approach is to minimize a cost function that measures the difference between forward solutions of a model and observations. The most direct, brute force method is sometimes called ''iterated solution of the forward problem:'' one minimizes the cost function through an exhaustive search. While this method can be used for multi-parameter inversions the computational burden can be quite large. For instance, if there are 10 parameter values each sampled at 20 values, then the search involves 10^{20} runs of the forward problem. Hence, nonlinear methods involving a more strategic search such as simulated annealing,[9] genetic algorithms[10] and others[11] have emerged as more practical alternatives. In

[a)]Portions of this work were presented in ''Adjoint-assisted inversion for shallow water environmental parameters,'' Proceedings of the Conference on Acoustic Variability, 16–20 September 2002, Lerici, La Specia, Italy, Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance, edited by N. G. Pace and F. B. Jensen (Kluwer, Dordrecht, 2002). Portions of this work were also presented at the 142nd Acoustical Society Meeting, 3–7 December 2001.

[b)]Electronic address: paul.hursky@saic.com

all of these forward solution techniques, the cost of running the forward model becomes the central factor. If the problem is only mildly nonlinear, a first attempt at decreasing this cost is to consider perturbations from known solutions so that the full cost of a massive forward search is avoided. Linear approximations to the sound pressure equation (as a function of medium properties) have been used in acoustic tomography.[12,13] Even with this approach, computing the perturbations over the full multi-parameter space remains the dominant, limiting factor.

Solutions to inverse problems in ocean acoustics usually include *a priori* information, such as gross knowledge of the ocean environment. For example, the sound speed is typically close to 1500 m/s. In some cases the starting guess is close enough to the true solution that we are within the multi-dimensional bowl of the cost function that contains the global minimum (unfortunately, it is rarely close enough for a linear inverse to work). When this occurs, it is useful to know the gradient of the cost function with respect to the large number of unknown parameters so as to guide us to the minimum. This gradient can be obtained by operating on the data-model misfit by the adjoint of the forward problem operator.[1,6] When the number of unknown parameters is much greater than the number of observations, the adjoint approach can be a less costly method of calculating the gradient than running the forward model for each variation in the unknown parameters.

For completeness, we mention that the derivative of a function mapping one normed vector space to another (e.g., sound speed versus range and depth to complex pressure amplitude across an array at a given frequency) is called a Fréchet derivative,[14] often mentioned as the key component of the inverse procedure. The adjoint method provides a procedure to compute the Fréchet derivative.[1,6,15] Adjoint methods also have an extensive earlier history in control theory, where they are derived from the Pontryagin principle.[14,16–18] Although we pursue a different approach in this paper, the adjoint operator can be derived by applying variational techniques to minimize an objective function that includes the data-model mismatch and the differential equations that describe the forward model being used, using Lagrange multipliers to set the constraint that the dynamical model is perfect. This produces the Euler–Lagrange equations, which are integrated by parts to reveal the adjoint operator and the boundary conditions that it must satisfy.[1,6,7,19] Adjoint methods have also been compared to the filtered back-propagation technique in diffraction tomography.[20]

The adjoint operator is formulated for a linearized forward model, so its application is limited to mildly nonlinear problems. This scope of problems is similar to those to which the Born approximation can be applied: a zeroth-order forward modeling calculation is used to calculate first-order perturbations to the medium, so the actual perturbations (being solved for) should not be so great or many that they drastically change the character of the zeroth-order field. To a certain extent this limitation can be overcome by iterating the approach, as we will show in Sec. III B.

This paper demonstrates inversion for sound speed fluctuations in the water column. Medium fluctuations and the variability they cause is a topic that is of much current interest, since variability introduces uncertainty into predictions of naval system performance. There is much ongoing work[21–24] addressing how medium fluctuations impact our ability to predict various processes in the ocean.

Although adjoint modeling is a concept that can be applied to all the standard acoustic models, the parabolic equation has a form that best matches the development of adjoint techniques in physical oceanography. The reason is that the PE marching solution in range is analogous to the marching solution in time of ocean circulation models. In addition, the parabolic equation can model range-dependent environments, where adjoint techniques provide the most payoff (many more unknowns than measurements). Section II reviews the parabolic equation, including its tangent linear and adjoint models, and derives an iterative procedure for solving inverse problems (Sec. II C contrasts the generalized inverse solution with solutions based on the adjoint). Section III presents results of testing the inversion procedure on simulated data based on sound speed profiles measured during the INTIMATE 96 experiment, during which internal tides were observed. Section III discusses the linearity of the PE in the particular experiment configuration used, demonstrates both range-independent and range-dependent inversions, and shows how the adjoint model can be used to calculate confidence bounds for estimated parameters.

II. THE PARABOLIC EQUATION FORWARD MODEL, ITS TANGENT LINEAR MODEL, AND ITS ADJOINT

The oceanographic community is quite actively developing adjoint techniques, so the oceanographic literature provides a useful starting point for formulating an acoustic adjoint.[4,5,19] The common oceanographic problem involves a marching forward in time of ocean temperature, currents, etc. through an ocean circulation model. One then observes the ocean conditions and seeks to understand how errors in initializing the ocean model, or in forcing it, caused those observation errors. The adjoint model directly back-propagates the data-model misfits, assuming there are no modeling errors, to provide corrections to the initial conditions and/or forcing.

In part, seeking a good acoustic analog of this process, we have in our initial work used a parabolic equation (PE) model which marches a starter acoustic pressure field forward in range from the source. The ocean and subbottom sound speed profile may be viewed as a forcing function and we will develop an adjoint model that back-propagates data-model misfits in the pressure field, thereby providing corrections to the sound speed profile that reduce these errors.

In using the PE to model acoustic propagation, modeling error is of much less concern than in ocean circulation models (which remain a work in progress), because the PE model is known to produce very accurate pressure predictions, whose error is very small compared to the changes in pressure caused by the sound speed profile changes that we are inverting for. As a result, we do not explicitly address modeling error in our inversion cost functions.

A. Perturbation of parabolic equation to produce tangent linear model

The first necessary ingredient for an adjoint model is the so-called tangent linear model. This is simply a linearized version of the PE that estimates (linearly) how perturbations in the sound speed profile map into perturbations in the pressure field. The standard homogeneous PE[25] (with no source in the medium) is

$$2ik_0\frac{\partial p}{\partial r}+\frac{\partial^2 p}{\partial z^2}+k_0^2[\,\gamma(r,z)-1\,]p=0, \tag{1}$$

where $\gamma(r,z)$ is the index of refraction squared, $c_0^2/c^2(r,z)$. Expanding this equation in terms of perturbations in pressure p and index of refraction squared γ to first order in ε,

$$\begin{aligned}\gamma&=\gamma+\varepsilon\tilde{\gamma},\\ p&=p+\varepsilon\tilde{p}+\cdots,\end{aligned} \tag{2}$$

yields

$$2ik_0\frac{\partial\tilde{p}}{\partial r}+\frac{\partial^2\tilde{p}}{\partial z^2}+k_0^2[\,\gamma(r,z)-1\,]\tilde{p}=-k_0^2\tilde{\gamma}(r,z)p.$$

We recognize that typical acoustic inversion problems will rarely be strictly linear. Therefore, in subsequent sections we will show an iterative method that relinearizes about its last estimate and can find its way to the solution despite mild nonlinearities.

B. Discretized PE and its tangent linear model

At this point, Eq. (2) can be discretized to implement a tangent linear model for the PE, from which an adjoint model can be constructed. However, it may be prudent to first derive a discretized PE—deriving an analytical adjoint and then discretizing it can introduce errors that break adjoint symmetry and result in suboptimal inversions.[19] Therefore, in this section, we start with a discretized PE (the implicit finite differences PE[25]) and derive a tangent linear model from its first-order terms.

1. Finite difference PE: Zeroth order

A finite difference version of Eq. (1) is

$$2ik_0\frac{\mathbf{p}^{n+1}-\mathbf{p}^n}{\Delta r}+\mathbf{A}\frac{\mathbf{p}^{n+1}+\mathbf{p}^n}{2}=0, \tag{3}$$

where $\mathbf{p}^n$ is the vector of pressures sampled in depth at range n, Δr is the size of the range step, and matrix $\mathbf{A}$ is the depth operator corresponding to the terms $\partial^2/\partial z^2+k_0^2[\,\gamma(r,z)-1\,]p$ in Eq. (1). Matrix $\mathbf{A}$ is

$$\begin{bmatrix} d_1 & e_2 & & & & & \\ e_2 & d_2 & e_3 & & & & \\ & e_3 & d_3 & & & & \\ & & & \ddots & & & \\ & & & & d_{N-2} & e_{N-1} & \\ & & & & e_{N-1} & d_{N-1} & e_N \\ & & & & & e_N & d_N \end{bmatrix}, \tag{4}$$

$$\begin{aligned} d_j&=-\frac{2}{h^2}+k_0^2[\,\gamma(r,z_j)-1\,],\\ e_j&=\frac{1}{h^2},\end{aligned} \tag{5}$$

where h is the depth increment (we have used a constant depth increment). Collecting terms

$$\left(\frac{2ik_0}{\Delta r}\mathbf{I}+\frac{\mathbf{A}}{2}\right)\mathbf{p}_{n+1}=\left(\frac{2ik_0}{\Delta r}\mathbf{I}-\frac{\mathbf{A}}{2}\right)\mathbf{p}_n, \tag{6}$$

which can be rewritten as

$$\mathbf{p}_{n+1}=\mathbf{F}\mathbf{p}_n, \tag{7}$$

where

$$\mathbf{F}=\left(\frac{2ik_0}{\Delta r}\mathbf{I}+\frac{\mathbf{A}}{2}\right)^{-1}\left(\frac{2ik_0}{\Delta r}\mathbf{I}-\frac{\mathbf{A}}{2}\right)=\mathbf{B}^{-1}\mathbf{C}. \tag{8}$$

This is the familiar PE marching solution, where $\mathbf{F}$ contains the range-dependent index of refraction squared $\gamma(r,z)$.

2. PE tangent linear model from first-order finite difference PE

To construct the adjoint of the PE model, we must first construct a tangent linear model for it. A first-order perturbation analysis of Eq. (8) yields

$$(\mathbf{B}+\varepsilon\tilde{\mathbf{B}})(\mathbf{p}_{n+1}+\varepsilon\tilde{\mathbf{p}}_{n+1})=(\mathbf{C}+\varepsilon\tilde{\mathbf{C}})(\mathbf{p}_n+\varepsilon\tilde{\mathbf{p}}_n), \tag{9}$$

where $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ are of opposite sign, and are both diagonal, with the ith diagonal element having magnitude $(k_0^2/2)\tilde{\gamma}(r,z_i)$. The $\tilde{\gamma}(r,z_i)$ are the first-order medium properties we are inverting for, and are indexed on depth z_i and range r.

The terms of order ε produce

$$\mathbf{B}\tilde{\mathbf{p}}_{n+1}=\mathbf{C}\tilde{\mathbf{p}}_n-\tilde{\mathbf{B}}\mathbf{p}_{n+1}+\tilde{\mathbf{C}}\mathbf{p}_n. \tag{10}$$

Replacing matrices $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ with diagonal matrix $\mathbf{D}$ and setting the appropriate factors yields

$$\mathbf{B}\tilde{\mathbf{p}}_{n+1}=\mathbf{C}\tilde{\mathbf{p}}_n-\frac{k_0^2}{2}\mathbf{D}(\mathbf{p}_{n+1}+\mathbf{p}_n). \tag{11}$$

Swapping the diagonal elements of $\mathbf{D}$ [i.e., $\tilde{\gamma}(r,z_i)$] with the elements of the vector that $\mathbf{D}$ is multiplying (i.e., $\mathbf{p}_{n+1}+\mathbf{p}_n$),

$$\mathbf{B}\tilde{\mathbf{p}}_{n+1}=\mathbf{C}\tilde{\mathbf{p}}_n-\Gamma\tilde{\gamma}(\mathbf{z}), \tag{12}$$

with Γ a diagonal matrix whose diagonal elements consist of the vector $(k_0^2/2)(\mathbf{p}_{n+1}+\mathbf{p}_n)$. Multiplying both sides by $\mathbf{B}^{-1}$,

$$\tilde{\mathbf{p}}_{n+1}=\mathbf{F}\tilde{\mathbf{p}}_n-\mathbf{G}\tilde{\mathbf{u}}_n, \tag{13}$$

$$\mathbf{G}=\mathbf{B}^{-1}\Gamma, \tag{14}$$

where $\tilde{\mathbf{u}}_n$ is the vector of $\tilde{\gamma}(r_n,z)$ values sampled in depth at range r_n [i.e., the ith element of $\tilde{\mathbf{u}}_n$ is $\tilde{u}_n(i)=\tilde{\gamma}(r_n,z_i)$, the first-order perturbation to the index of refraction squared at range n and depth i]. Matrix $\mathbf{F}$ in Eq. (13) is identical to the $\mathbf{F}$ in Eq. (8) for the zeroth-order pressures. Equation (13) propagates the first-order pressure field $\tilde{\mathbf{p}}$ (the perturbation to

the zeroth-order pressure field $\mathbf{p}$), with the first-order medium perturbations $\tilde{\mathbf{u}}_n$ [i.e., $\tilde{\gamma}(r,z)$] acting as a forcing function. The perturbations $\tilde{\gamma}(z)$, and matrices $\mathbf{B}$ and $\mathbf{C}$, by which $\mathbf{F}$ and $\mathbf{G}$ are represented, can be range dependent. Equation (13) is the tangent linear model for our (otherwise nonlinear) PE model. It is important to keep in mind that the tangent linear model and the adjoint model derived from it in Sec. II D propagate quantities that are first-order perturbations.

C. How adjoint model arises in our PE-based inversion

In this section, we combine the sequence of PE marching steps into a single linear system, and show how the adjoint of this system back-propagates data-model misfit at the observation points back to the points in the medium where adjustments should be made. The purpose of this section is to illustrate what an adjoint model does in a very concrete way. We will actually use a more efficient adjoint formulation presented in the next section to solve inverse problems in later sections.

To simplify notation, we drop the tilde notation that was used in Sec. II B to indicate a first-order perturbation and assume all references to $\mathbf{p}$ and $\mathbf{u}$ below are to their first-order terms.

Equation (13) generates the following equations, one for each range step (the subscripts indicate the index of the range step):

$$\begin{aligned}
&\mathbf{p}_1 = \mathbf{F}_0\mathbf{p}_0 - \mathbf{G}_0\mathbf{u}_0, \qquad (15)\\
&\mathbf{p}_2 = \mathbf{F}_1\mathbf{F}_0\mathbf{p}_0 - \mathbf{F}_1\mathbf{G}_0\mathbf{u}_0 - \mathbf{G}_1\mathbf{u}_1,\\
&\vdots\\
&\mathbf{p}_N = \mathbf{F}_{N-1}\mathbf{F}_{N-2}\cdots\mathbf{F}_1\mathbf{F}_0\mathbf{p}_0\\
&\qquad - \sum_{n=0}^{N-1} \mathbf{F}_{N-1}\mathbf{F}_{N-2}\cdots\mathbf{F}_{n+1}\mathbf{G}_n\mathbf{u}_n .
\end{aligned}$$

We include vector $\mathbf{p}_0$, the correction to the presumed starter field, as an unknown here for completeness. We will only solve for medium properties in the sections that follow. Each pressure $\mathbf{p}_n$ is a function of all the medium property vectors $\mathbf{u}_i$, $i=1,\ldots,n$, leading up to it. A matrix representation of this summation is

$$[\mathbf{A}\ \ \mathbf{B}_0\ \ \mathbf{B}_1\ \ \cdots\ \ \mathbf{B}_{N-1}]\begin{bmatrix}\mathbf{p}_0\\ \mathbf{u}_0\\ \mathbf{u}_1\\ \vdots\\ \mathbf{u}_{N-1}\end{bmatrix} = \mathbf{p}_N, \qquad (16)$$

where the submatrix $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{F}_{N-1}\mathbf{F}_{N-2}\cdots\mathbf{F}_0, \qquad (17)$$

each submatrix $\mathbf{B}_n$ is

$$\mathbf{B}_n = -\mathbf{F}_{N-1}\mathbf{F}_{N-2}\cdots\mathbf{F}_{n+1}\mathbf{G}_n, \qquad (18)$$

and $\mathbf{p}_0$ and all the individual $\mathbf{u}_n$ vectors have been stacked one on top of each other in Eq. (16) to form the super vector $\mathbf{u}$. Equation (16) is a representation of the inverse problem in a familiar linear system form:

$$\mathbf{F}\mathbf{u} = \mathbf{p}_N, \qquad (19)$$

where $\mathbf{F}$ and $\mathbf{u}$ are composite matrices.

Equation (19) can be solved using either a generalized inverse of $\mathbf{F}$,

$$\hat{\mathbf{u}} = \mathbf{F}^{+}\mathbf{p}_N, \qquad (20)$$

or an iterative steepest descent technique,

$$\hat{\mathbf{u}}_i = \hat{\mathbf{u}}_{i-1} + \alpha\mathbf{F}^{*}(\mathbf{F}\mathbf{u}_{i-1} - \mathbf{p}_N), \qquad (21)$$

in which the second term on the right hand side is just the gradient of the squared data-model misfit, scaled by a step size parameter α. Note how the gradient is formed by the adjoint $\mathbf{F}^*$ operating on the data-model misfit $\mathbf{F}\mathbf{u}_{i-1} - \mathbf{p}_N$.

From Eq. (16), the conjugate transpose of $\mathbf{F}$ (the adjoint of our linearized PE) is

$$\mathbf{F}^{*} = \begin{bmatrix}\mathbf{A}^{*}\\ \mathbf{B}_0^{*}\\ \vdots\\ \mathbf{B}_{N-1}^{*}\end{bmatrix}, \qquad (22)$$

where the individual matrix

$$\mathbf{B}_n^{*} = -\mathbf{G}_n^{*}\mathbf{F}_{n+1}^{*}\cdots\mathbf{F}_{N-2}^{*}\mathbf{F}_{N-1}^{*} \qquad (23)$$

can be interpreted as the transfer matrix which back-propagates the data-model misfit (often identified as the adjoint field) from the receiver at range N to points back in the medium at range n. Matrix $\mathbf{F}_{N-1}^{*}$ steps from range N to $N-1$. Matrix $\mathbf{F}_{N-2}^{*}$ steps from range $N-1$ to $N-2$, and so on, until at range n, $\mathbf{G}_n^{*}$ transforms the adjoint field to a medium property correction at range step n (as we will show later). The adjoint of our forward model propagates information from the observations at the receiver array back to all the medium properties that we are inverting for (i.e., the $\mathbf{u}_n$ at all ranges n). Note that matrix $\mathbf{A}^*$ similarly propagates information all the way back to the source, where it can be used to correct the starter field (if this is treated as an unknown).

D. Fixed-point iteration to invert for medium properties

In this section, we present an alternative derivation of the adjoint model, using Lagrange multipliers. This yields a fixed-point iteration, which we relate to a steepest-descent iteration. To simplify the presentation, as in Sec. II C, we drop the tilde notation that was used in Sec. II B to indicate first-order perturbations and assume all references to $\mathbf{p}$ and $\mathbf{u}$ are to their first-order terms.

The first-order parabolic equation is

$$\mathbf{p}_{n+1} = \mathbf{F}_n\mathbf{p}_n - \mathbf{G}_n\mathbf{u}_n, \qquad (24)$$

where the unknown medium properties $\mathbf{u}_n$ act as forcing functions at each range step n of the marching solution [Eq. (24) is the same as Eq. (13), without the tilda notation]. Matrices $\mathbf{F}_n$ and $\mathbf{G}_n$ are functions only of the zeroth-order pressures and medium properties, and not their first-order perturbations, $\mathbf{p}_n$ and $\mathbf{u}_n$.

To solve for $\mathbf{u}_n$, we formulate a cost function $J(\mathbf{p},\mathbf{u},\lambda)$ to be minimized:

$$J(\mathbf{p},\mathbf{u},\lambda)=\frac{1}{2}(\mathbf{p}_N-\mathbf{m}_N)^*(\mathbf{p}_N-\mathbf{m}_N)+\sum_{n=1}^{N}\lambda_n^*(\mathbf{p}_n-\mathbf{F}_{n-1}\mathbf{p}_{n-1}+\mathbf{G}_{n-1}\mathbf{u}_{n-1})+\frac{\beta}{2}\sum_{n=0}^{N-1}\mathbf{u}_n^*\mathbf{u}_n. \tag{25}$$

The first term in J seeks to minimize the mismatch between the measured $\mathbf{m}_N$ and modeled $\mathbf{p}_N$ pressure increments, both at range index N. The data-model misfit in the first term will also contain measurement noise in $\mathbf{m}_N$, which we ignore, since we focus on the underdetermined case (which provides no redundant measurements that can be used to smooth the solution). Since we are dealing with first-order terms, $\mathbf{m}_N$ is the difference between the measured pressure and the zeroth-order pressure prediction. The modeled pressure $\mathbf{p}_N$ is calculated by propagating $\mathbf{p}_n$ from the source to the receiver using Eq. (24) [i.e., our tangent linear model $\{\mathbf{F}_n,\mathbf{G}_n\}$ with the estimated environmental parameters $\mathbf{u}_n$ as driving functions]. Given a solution for $\mathbf{u}_n$, the zeroth-order pressure plus the pressure correction $\mathbf{p}_N$ calculated using $\mathbf{u}_n$ should reproduce the measured pressure. The second term uses Lagrange multipliers λ_n (vectors at each range, sampled in depth) to enforce the hard constraint that the model is perfect (i.e., $\mathbf{p}_n$ and $\mathbf{u}_n$ are completely governed by the linearized forward model $\{\mathbf{F}_n,\mathbf{G}_n\}$). The third term is a regularizing term to minimize the amplitude of the medium property perturbations $\mathbf{u}_n$, with β a weighting term to regulate how much we allow the $\mathbf{u}_n$ to stray from our assumed zeroth-order values (later, we will set β to the ratio of the variances of the pressure measurements and the medium properties).

Admittedly, this is an unusual way to formulate an inverse problem. We have set up a large number of unknowns in all the intermediate $\mathbf{p}_n$, in addition to the already large number of unknowns $\mathbf{u}_n$. We will show that minimizing the cost function in Eq. (25) is a more direct way of inverting for the $\mathbf{u}_n$ than repeatedly running the forward model to explore the surface J as a function of $\mathbf{u}_n$.

Note that J is a function of $\mathbf{u}_n$ at all ranges and depths, so its minimization can be used to resolve range-dependent features.

The partial derivatives of $J(\mathbf{p},\mathbf{u},\lambda)$ are

$$\frac{\partial J}{\partial \mathbf{p}_N}=\mathbf{p}_N-\mathbf{m}_N+\lambda_N, \tag{26}$$

$$\frac{\partial J}{\partial \mathbf{p}_n}=\lambda_n-\mathbf{F}_n^*\lambda_{n+1}, \tag{27}$$

$$\frac{\partial J}{\partial \mathbf{u}_n}=\beta\mathbf{u}_n+\mathbf{G}_n^*\lambda_{n+1}, \tag{28}$$

$$\frac{\partial J}{\partial \boldsymbol{\lambda}_n}=\mathbf{p}_n-\mathbf{F}_{n-1}\mathbf{p}_{n-1}+\mathbf{G}_{n-1}\mathbf{u}_{n-1}. \tag{29}$$

Setting these partial derivatives to zero yields a system of equations which are solved to invert for the sound speed profile. Much of the adjoint literature describes iterative processes to solve these equations. It should be emphasized that the iteration is not something that directly follows in logical sequence but is simply a proposed approach to solving this large system. Actually, it is just a so-called fixed-point or functional iteration (also called a Picard iteration).

To explain the iteration in those terms, the derivatives in Eqs. (26)–(28) are set to zero and rearranged to yield

$$\mathbf{u}_n=f(\mathbf{u}_n) \tag{30}$$

with

$$f(\mathbf{u}_n)=-\frac{1}{\beta}\mathbf{G}^*\lambda_{n+1}. \tag{31}$$

The term $\mathbf{G}_n^*\lambda_{n+1}$ comes from the back-propagation of the λ_N to λ_n, with λ_N a function of $\mathbf{p}_N$, and $\mathbf{p}_N$ itself calculated from $\mathbf{u}_n$. The idea of a fixed-point iteration is to pick a starting guess, $\mathbf{u}^0$, and apply a simple recursion:

$$\mathbf{u}^{i+1}=f(\mathbf{u}^i). \tag{32}$$

Whether the fixed-point iteration converges or not is often unclear. However, if it does, then we have solved for the solution of our original equation. In fact, part of the art of constructing a good fixed-point iteration is to rearrange the original equation into a form where the iteration converges as rapidly as possible. For instance, we can obtain a broader class of iterations by multiplying Eq. (30) by α and adding $\mathbf{u}_n$ on both sides,

$$\mathbf{u}_n+\alpha\mathbf{u}_n=\mathbf{u}_n+\alpha f(\mathbf{u}_n), \tag{33}$$

which is rearranged as

$$\mathbf{u}_n=\mathbf{u}_n-\alpha(\mathbf{u}_n-f(\mathbf{u}_n)). \tag{34}$$

Now, α provides a parameter that we can adjust to optimize the rate of convergence. To summarize the fixed-point iteration in terms of the variables in our problem, we have the following recursion:

$$\mathbf{p}_{n+1}^{i+1}=\mathbf{F}_n\mathbf{p}_n^{i+1}-\mathbf{G}_n\mathbf{u}_n^i, \tag{35}$$

$$\lambda_N^{i+1}=\mathbf{m}_N-\mathbf{p}_N^{i+1}, \tag{36}$$

$$\lambda_n^{i+1}=\mathbf{F}_n^*\lambda_{n+1}^{i+1}, \tag{37}$$

$$\mathbf{u}_n^{i+1}=\mathbf{u}_n^i-\alpha\left(\mathbf{u}_n^i+\frac{\mathbf{G}_n^*\lambda_{n+1}^{i+1}}{\beta}\right). \tag{38}$$

In words, we calculate $\mathbf{p}_n^{i+1}$ using our forward model to propagate pressures from the source to the receiver based on our current estimate of $\mathbf{u}_n^i$ in Eq. (35), initialize λ_N^{i+1} at the receiver to the data-model misfit [Eq. (36)], propagate the λ_n^{i+1} from the receiver to the source in Eq. (37), and calculate the updated first-order medium properties $\mathbf{u}_n^{i+1}$ from the updated λ_n^{i+1} at all range indexes n in Eq. (38).

Equation (37) represents the *adjoint model*. The λ_n's form the *adjoint field* propagated by the adjoint model $\mathbf{F}_n^*$. Equation (36) initializes the adjoint field λ_N to the data-model misfit at the receiver. Note the complementary relationship between the pressures $\mathbf{p}_n$, being propagated from

the source to the receiver by our forward model $\mathbf{F}_n$, and the Lagrange multipliers λ_n, being propagated by the adjoint model $\mathbf{F}_n^*$ backward from the receiver to the source.

A compelling interpretation of the adjoint is that the observed field is a superposition of a baseline field due to the presumed medium and a perturbed field due to the unknown medium perturbations (which we invert for). The adjoint model back-propagates (time-reverses) the perturbed field to the unknown medium perturbations, which are viewed as sources of diffraction.[1]

An important insight about this fixed-point iteration is that it is also equivalent to a descent technique. To see this, observe that the second term in Eq. (34) [also Eq. (38)] is proportional to the gradient of J with respect to $\mathbf{u}$ as given in Eq. (28). Thus, our fixed-point iteration is equivalent to an iteration

$$\mathbf{u}^{i+1} = \mathbf{u}^i - \frac{\alpha}{\beta}\frac{\partial J}{\partial \mathbf{u}}. \tag{39}$$

In this form, we see that the parameter α plays the usual role in a steepest descent method of controlling the step size in the descent direction. It is natural in this descent framework to consider all the standard extensions such as conjugate gradient techniques.[26] We do not discuss these options further here, but simply note they are part of the standard set of techniques used to improve the convergence of the adjoint iteration.

The fixed-point iteration equations derived above are in terms of $\mathbf{F}_n$ and $\mathbf{G}_n$ which are functions of the zeroth-order medium properties and pressures. As a result, the iterative process can periodically be relinearized about the current estimates for $\mathbf{u}_n$ in order to update the tangent linear model. This enables the process to handle mildly nonlinear problems.

III. INVERTING FOR INTIMATE 96 INTERNAL TIDES (SIMULATED RESULTS)

In this section, the adjoint method described in Sec. II D is tested on synthetic acoustic data calculated for a series of sound speed profiles measured during the INTIMATE 96 experiment[27] during which internal tides were observed. Figure 1 shows the entire sequence of measured sound speed profiles, in the order that they were measured (the sampling in time was not uniform, so indexes rather than time units are shown on the horizontal axis). Figure 2 shows the entire set of measured sound speed profiles and the source-receiver configuration used in our simulation. These profiles were measured over roughly 60 h. The experiment configuration was fixed-fixed. The line joining the source and receiver was perpendicular to the movement of internal tides during part of the actual experiment. Such a configuration would suggest a *range-independent* sound speed field in the plane joining the source and receiver.

An important step when applying adjoint methods is to ensure that the problem has been linearized, so Sec. III A compares the pressures calculated by the original PE and the linearized PE models, for the configuration to be used to demonstrate the inversion. Section III B applies the iterative

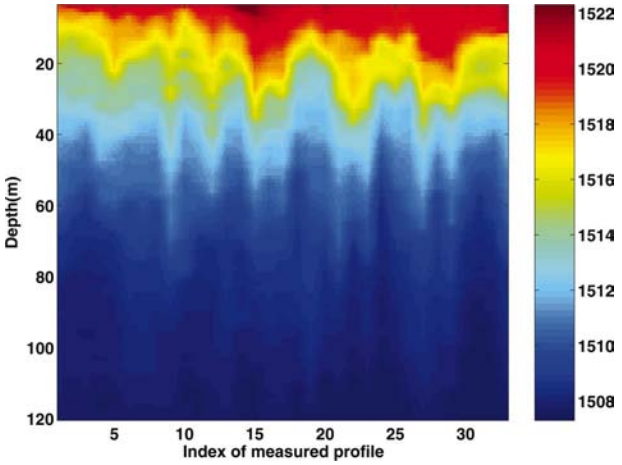


FIG. 1. Sound speed profiles measured during INTIMATE 96 over a time interval of roughly 3 days. The sound speed profiles were measured at only approximately regular time intervals, so the horizontal axis shows the index of the measurements. The image scale shows sound speed measured in m/s.

adjoint technique described in Sec. II to a series of *range-independent* sound speed fields. Section III C shows that our adjoint implementation reproduces the tangent linear model (using far fewer modeling runs than if the forward model alone were used). Section III D shows how a generalized inverse of the linearized forward model can be used as an alternative to our iterative adjoint inversion, and also how to calculate error bars for these inversion results. Section III E applies the iterative adjoint inversion to a range-dependent inversion problem.

The PE model was used to synthesize acoustic pressures to serve as ''measured'' data in our simulations. Pressures were calculated at 400 Hz in Sec. III B for the range-independent case and for a range of frequencies from 100 to 400 Hz in Sec. III E for the range-dependent cases. The receiver was a fully spanning vertical line array at a 2-km distance from the source. The source depth was at 50-m depth in Sec. III B for the range-independent case. Sources spanned 5 to 90 m for the range-dependent cases in Sec.

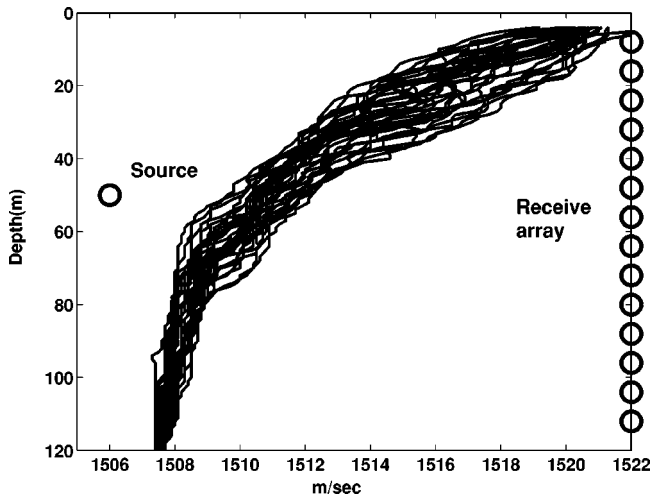


FIG. 2. Sound speed profiles measured during INTIMATE 96 and experimental configuration for which acoustic data was synthesized. All inversions were performed at a frequency of 400 Hz and a range of 2 km.

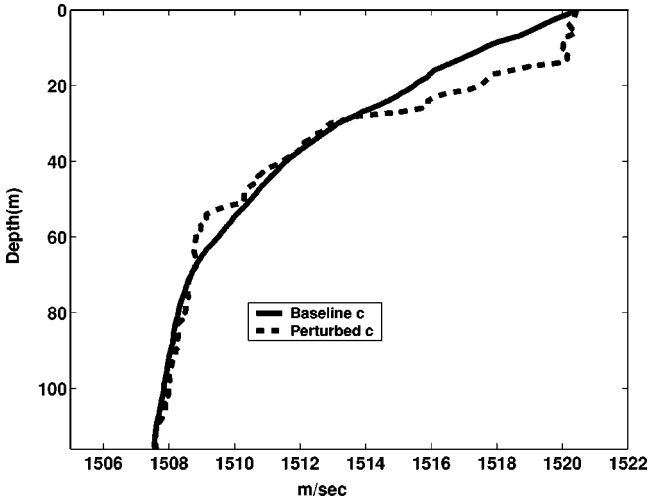


FIG. 3. Perturbed and baseline sound speed profiles used to test linearity of INTIMATE 96 configuration.

III E, with each source an independent omni-directional point source and not part of a transmitting array.

A. Testing our tangent linear model to verify we are in linear regime

Prior to using the adjoint method for inversion it is important to assess whether the tangent linear model and its adjoint are reasonable approximations to the full PE model in this experimental configuration. Using the same experimental configuration as in our inversions, we compare perturbations produced by our tangent linear model with perturbations predicted by the original (nonlinear) PE. Figure 3 shows two sound speed profiles, $c(z)$ to be used as a baseline, and $c(z)+\delta c(z)$ as a perturbed profile. The baseline profile is the mean profile calculated from the entire set of INTIMATE 96 sound speed measurements. The perturbed profile is one chosen from the measured profiles for its large deviation from the mean profile. Figure 4 shows the pressure field produced by the original PE model operating on the baseline sound speed profile ($P(c)$, where P represents the

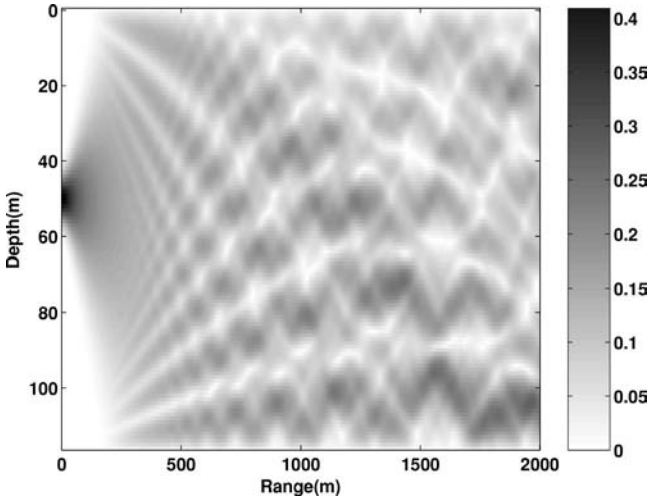


FIG. 4. Baseline pressure magnitude at 2 km, source depth of 50 m, and at 400 Hz.

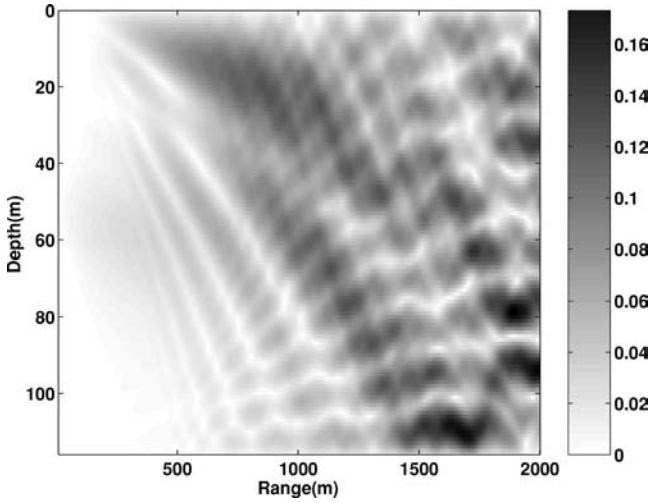


FIG. 5. Magnitude difference between perturbed and baseline pressures.

original PE). Figure 5 shows the magnitude difference between perturbed and baseline pressures calculated by the original PE [i.e., $P(c+\delta c)-P(c)$]. Figure 6 shows the magnitude difference between both sets of perturbed pressures, those calculated by the original PE and those calculated by its tangent linear model [i.e., $P(c+\delta c)-P(c)-\mathbf{F}(c)\delta c$, where $\mathbf{F}(c)$ represents the tangent linear model, linearized about c]. Comparing the images in Fig. 5 and Fig. 6 (whose color scales are the same) shows that the perturbation in pressures due to the change in medium properties (in Fig. 5) is always larger than the size of the error between the original PE model and its tangent linear model (in Fig. 6). This shows that the problem is not too nonlinear (our iterative process will relinearize after each iteration), and that our configuration is reasonable for testing our iterative adjoint process.

As we have shown in Secs. II C and II D, the adjoint process is a local optimization, so if we are outside the solution's ''basin of attraction,'' there is no mechanism for escaping a spurious local minimum. This can happen if the initial guess is too far from the correct solution, or if the linear approximation being used by the tangent linear model

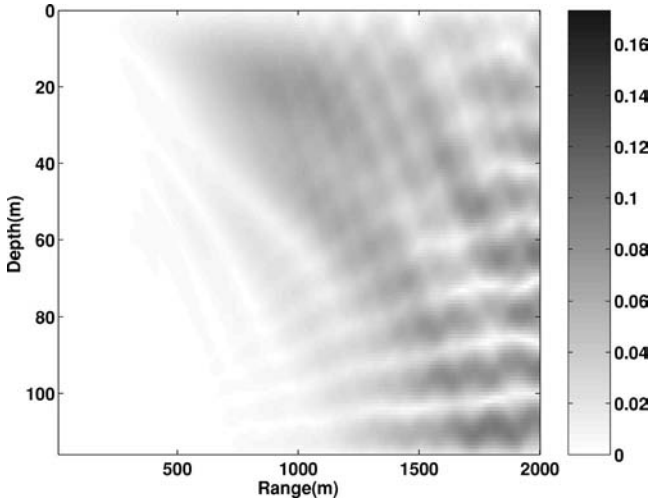


FIG. 6. Magnitude difference between perturbed and tangent linear pressures (would be zero if tangent linear model perfectly predicted the perturbed pressure). Compare to differences shown in Fig. 5.

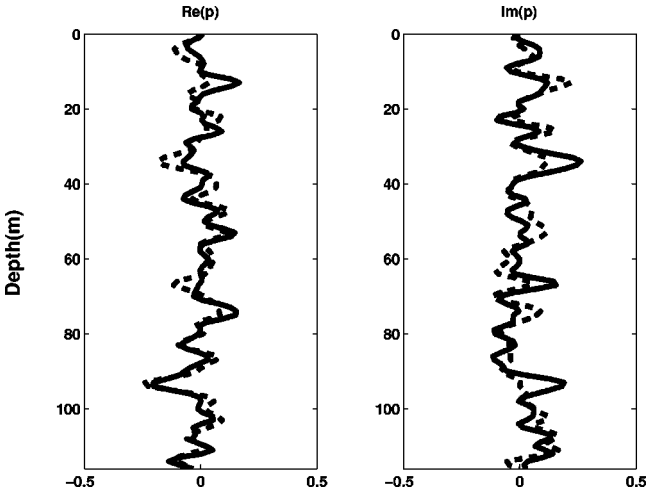


FIG. 7. ''Measured'' pressure was synthesized using PE with true sound speed profile (solid line). Modeled pressure was synthesized using the PE model with mean sound speed profile (dashed line). Real components are shown in left plot. Imaginary components are shown in right plot.

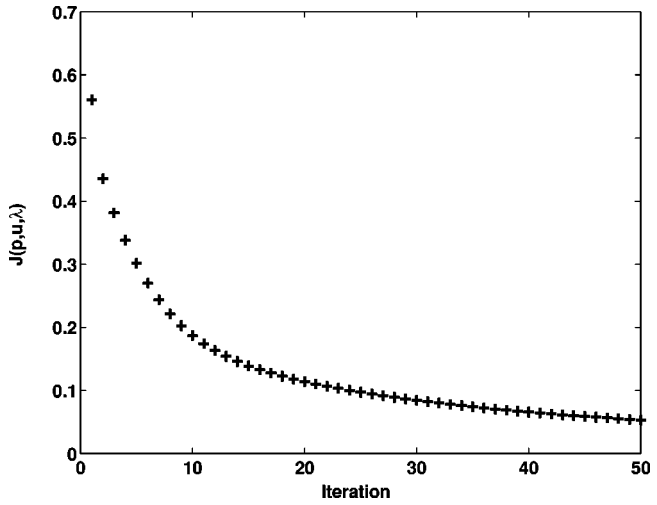


FIG. 8. Evolution of objective function produced by iterative adjoint process.

(and its adjoint) fails to reproduce the true cost function. Because it relinearizes about its last estimate and can follow a nonlinear basin of attraction if started reasonably close to the true solution, our iterative adjoint process can handle slightly nonlinear problems.

B. Adjoint inversions using fixed-point iteration (range-independent case)

In this section, the iterative adjoint inversion process discussed in Sec. II D is tested on *synthetic* acoustic data to assess the feasibility of inverting for sound speed profiles that indicate the internal tides present during the INTIMATE 96 experiment. For each measured profile from INTIMATE 96, an acoustic pressure vector is calculated by the PE model to serve as a ''measured'' pressure vector for our inversions. Each of these ''measured'' pressures is inverted to estimate the sound speed profile used to synthesize it. The mean of the entire set of profiles serves as the initial guess in each inversion.

Before showing the results on the entire sequence of profiles, we show the details of an inversion on a single profile, profile 28, selected for its comparatively large deviation from the mean. Figure 7 shows the ''measured'' pressure (calculated using profile 28) and the modeled pressure upon initialization (calculated for the mean profile, used to initialize our iterative process). A single iteration consists of using the forward model to calculate a modeled pressure, then using the adjoint model to calculate a corrected sound speed profile. Figure 8 shows cost function values [see Eq. (25)] at iterations 1–50. Figure 9 shows the modeled pressures and the measured pressures after 50 iterations. Figure 10 shows a comparison between the true and estimated sound speed profile perturbations (i.e., after 50 iterations).

The inversion shown in detail for profile 28 (in Figs. 7–10) was repeated at each of the profiles measured during the INTIMATE 96 experiment. Figure 11 shows the inversion results for the entire set of INTIMATE 96 sound speed profiles. The sound speed profiles were measured at only approximately regular time intervals, so the horizontal axes show the indexes of the measurements. The upper left plot shows the ''measured'' profiles (the true profiles, in our simulation). The lower left plot shows the estimated profiles after 50 iterations at each profile. The upper right plot shows the ''true'' profiles minus the mean profile. These are the first order quantities that are being estimated. The lower right plot shows the difference between the ''true'' and estimated sound speed profiles. Both the upper right and lower right plots have the same scale, so that the estimation errors can be compared with the quantities being estimated. The estimation errors are significantly smaller. This experiment was repeated with only four receiver elements (using the same configuration as the INTIMATE 96 experiment) and at a number of signal-to-noise ratios (additive white gaussian noise was added at each of the four receivers). Similar results to those shown in Fig. 11 were obtained at SNRs greater than 10 dB.

C. Using adjoint model to construct tangent linear model (range-independent case)

In this section, partly to assess the validity of our adjoint implementation, we calculate the tangent linear model using

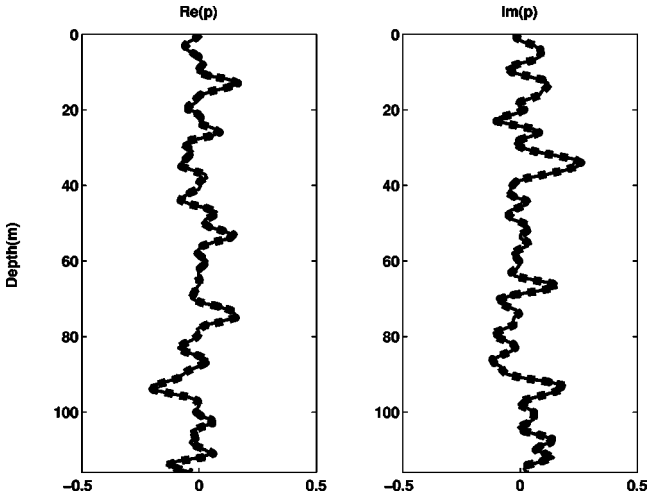


FIG. 9. Measured and final estimated pressure magnitudes.

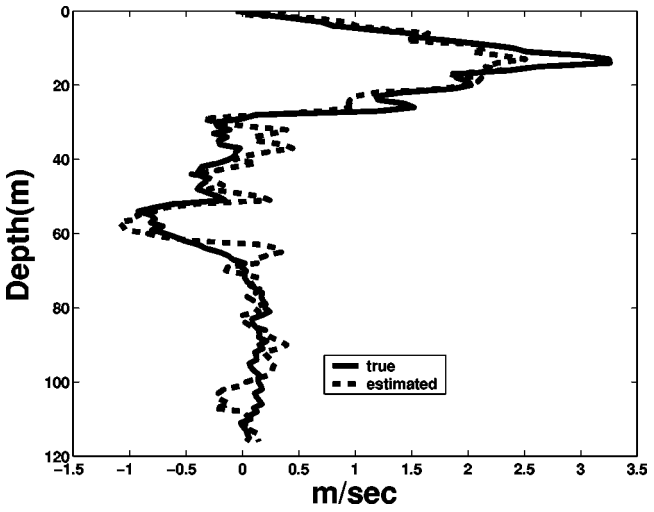


FIG. 10. True and final estimated sound speed profile perturbations.

our adjoint model and compare it with a tangent linear model calculated in a brute force fashion using the original PE. If our tangent linear model **F** is **M** by **N** (**M** observations, **N** waveguide parameters), we will use **M** runs of our adjoint model to calculate the **M** rows of **F** instead of using **N** runs of our forward model to calculate the **N** columns of **F**. This comparison is made at one of the profiles estimated in Sec. III B (profile 28). With a much larger number of medium properties than measurements, the adjoint model provides a much less computationally demanding means of calculating the tangent linear model. As we will show in Sec. III D, such a tangent linear model can be used to form a generalized inverse solution to the inverse problem (an alternative to the iterative process we have demonstrated above) and to calculate error bars around the inverse problem solution.

The tangent linear model is the matrix **F** whose element F_{ij} is the ith observed pressure due to a unit perturbation of the jth index of refraction squared value. Thus, if we use the configuration of source and a vertical line receive array (as in our simulated example), a column of **F** can be calculated by two runs of our unperturbed PE model:

$$\mathbf{f}_i = P(\mathbf{u}+\tilde{\mathbf{u}}_i) - P(\mathbf{u}), \tag{40}$$

where $\mathbf{f}_i$ represents the ith column of **F**, and $P(\mathbf{u})$ represented the pressures calculated at the receive array elements by the original PE model using medium properties **u**. The perturbation in **u** is $\tilde{\mathbf{u}}_i = \alpha \mathbf{e}_i$, where $\mathbf{e}_i$ is the ith elementary basis vector (having a one at its ith element and zeros at all other elements), and α is a scaling factor small enough to keep the model linear. Each $\mathbf{f}_i$ is a vector of pressure perturbations at the vertical line receive array elements due to the perturbation of **u** at its ith element. The matrix **F** is rescaled by dividing it by α, to compensate for the earlier scaling of the index of refraction perturbations $\tilde{\mathbf{u}}_i$. We will call the matrix calculated using the forward model $\mathbf{F}_u^{\text{forward}}$.

For the adjoint approach, we run the adjoint model to calculate the sound speed correction that would reproduce pressure perturbations at each individual receiver,

$$\tilde{\mathbf{u}}_j = \mathbf{F}^*(\tilde{\mathbf{p}}_j), \tag{41}$$

where $\mathbf{F}^*$ is the adjoint model [see Eqs. (37) and (38)] and $\tilde{\mathbf{p}}_j = \alpha \mathbf{e}_j$. The conjugates of vectors $\tilde{\mathbf{u}}_j$ are put into the rows of a new matrix, $\mathbf{F}_u^{\text{adjoint}}$. The $\tilde{\mathbf{u}}_j$ are often called the representer adjoints, which must be calculated prior to calculating the representers themselves in a process which converts the inverse problem from a two-point boundary value problem to a linear algebraic problem in which the solution is expressed as a linear combination of representers.[19]

We compared the left eigenvectors and singular values (calculated by a singular value decomposition) of $\mathbf{F}_u^{\text{forward}}$ with $\mathbf{F}_u^{\text{adjoint}}$ to assess how well the adjoint model reproduced the tangent linear model calculated by the unperturbed PE. At the significant eigenvectors, the adjoint model was able to reproduce the eigenstructure of the $\mathbf{F}_u^{\text{forward}}$ matrix.[28]

D. Generalized inverse solution and posterior covariances (range-independent case)

In this section, we use the representer adjoints calculated in the previous section to calculate a generalized inverse solution, as discussed in Sec. II C. We show this calculation as a final correction to the estimate produced by our iterative adjoint process in Sec. III B, but it could in principle replace the iterative adjoint-based inversion. We also show how the representer adjoints can be used to calculate the *posteriori* covariances of our sound speed estimates, given prior covariances for our pressure measurements and medium properties. All the calculations in this section (as in the previous section) are performed at profile 28, estimated in Sec. III B.

Prior knowledge of measurement and medium property error statistics is typically incorporated into the cost function we seek to minimize using covariance matrices $\mathbf{C}_p$ (for pressure measurement errors **p**) and $\mathbf{C}_u$ (for medium properties **u**). These covariance matrices map the data vector space and model vector space to their dual spaces, enabling a consistent formulation of the optimization in terms of a gradient (which is an element of the dual space).[29] Although incorporating more realistic covariances is usually necessary in practical applications, we will simplify the material that follows by assuming the covariances are scaled identity matrices:

$$\mathbf{C}_p = \sigma_p^2 \mathbf{I}, \quad \mathbf{C}_u = \sigma_u^2 \mathbf{I}, \tag{42}$$

and use $\beta = \sigma_p^2/\sigma_u^2$ to represent our prior expectation of the relative accuracy of our measurements compared with that of our baseline medium properties. For high SNR, we set β low (σ_p small compared with σ_u), allowing our **u** estimates to stray further from our prior expectations (the zeroth-order medium properties). Note that although we have not explicitly included measurement errors in Eq. (25), it is addressed implicitly by β, (and in general by $\mathbf{C}_p$ and $\mathbf{C}_u$ when they are not scaled identity matrices).

Defining a cost function $J(\mathbf{u})$ in terms of the data-model misfit and norm of **u** (to penalize large **u**),

$$J(\mathbf{u}) = \tfrac{1}{2}(\mathbf{F}\mathbf{u}-\mathbf{m}_N)^*\mathbf{C}_p^{-1}(\mathbf{F}\mathbf{u}-\mathbf{m}_N) + \tfrac{1}{2}\mathbf{u}^*\mathbf{C}_u^{-1}\mathbf{u}, \tag{43}$$

setting the gradient of $J(\mathbf{u})$ to zero, and solving for **u** produces

$$\mathbf{u} = (\mathbf{F}^*\mathbf{C}_p^{-1}\mathbf{F} + \mathbf{C}_u^{-1})^{-1}\mathbf{F}^*\mathbf{C}_p^{-1}\mathbf{m}_N, \tag{44}$$

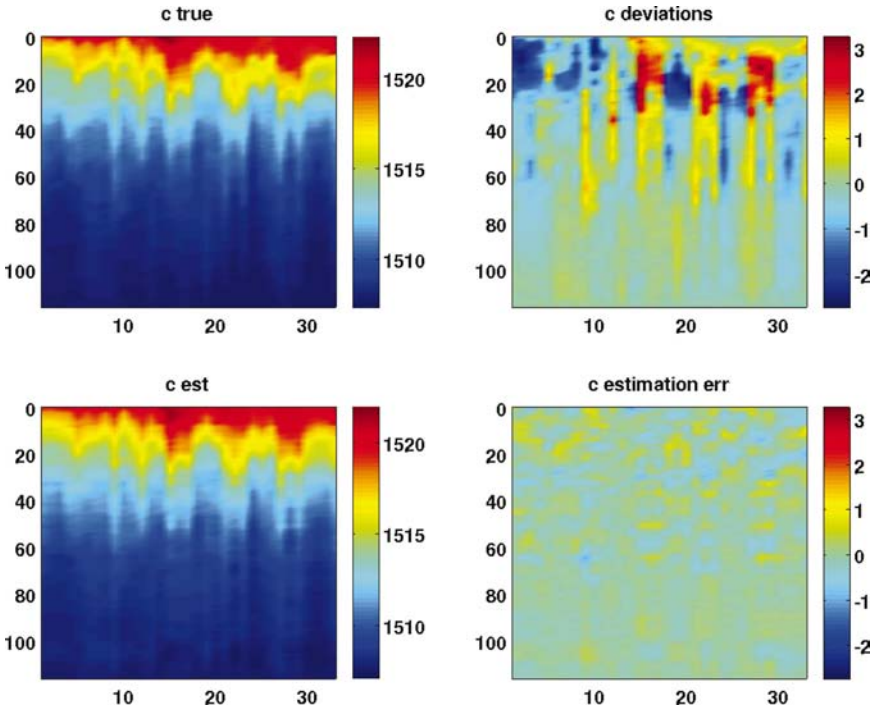


FIG. 11. Tracking internal tides over several diurnal cycles during INTIMATE 96 experiment. Horizontal axes are the sample numbers of the measured profiles. Note the different image scales in the left and right columns (the same image scale is used in each column so that the upper and lower plots can be compared). All image scales are sound speeds measured in m/s.

in which the matrix inverse is performed in the space of model parameters (analogous to a least-squares solution if the regularizing term $\mathbf{C}_u^{-1}\mathbf{u}$ is omitted). An alternative form,

$$\mathbf{u}=\mathbf{C}_u\mathbf{F}^*(\mathbf{F}\mathbf{C}_u\mathbf{F}^*+\mathbf{C}_p)^{-1}\mathbf{m}_N, \tag{45}$$

enables the matrix inverse to be performed in the data space (analogous to a minimum-norm solution). These solutions both involve the calculation of a potentially expensive generalized inverse—the choice can be made by which of the two spaces, model or data, is smaller.

The inversions in Sec. III B produced estimates of $\mathbf{u}$ and residual errors $\partial\mathbf{p}$ (data-model misfit after 50 iterations). Setting $\mathbf{C}_p$ and $\mathbf{C}_u$ to be scaled identity matrices and β to 10^{-3} in Eq. (45), we substitute $\partial\mathbf{p}$ for $\mathbf{m}_N$ and calculate $\partial\mathbf{u}$, a final correction to our previous estimate $\mathbf{u}$. Figure 12 shows the adjoint solution from the previous section, the true solution, and the solution refined by the generalized inverse of $\mathbf{F}_\mathbf{u}$. At almost all depths, the generalized inverse has improved the adjoint estimates.

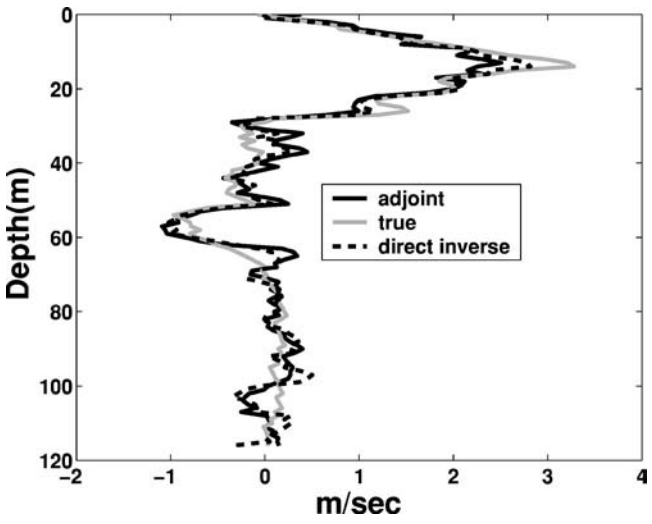


FIG. 12. Gray line shows sound speed estimate produced by iterative adjoint process. Black line shows the true sound speed. Dashed line shows the result of using generalized inverse to further reduce the residual errors from our iterative process.

To calculate error bars for our estimates of $\mathbf{u}$, we assume uncorrelated Gaussian errors and that our estimate is close to the true sound speed (so that a linear assumption holds). When this is the case, the *posterior* error covariance matrix (of the medium property estimation errors) is defined in terms of a prior pressure measurement covariance $\mathbf{C}_p$ and a prior medium property covariance matrix $\mathbf{C}_u$ as

$$\mathbf{R}_u=(\mathbf{F}_u^*\mathbf{C}_p^{-1}\mathbf{F}_u+\mathbf{C}_u^{-1})^{-1} \tag{46}$$

or, in the space of measurements,

$$\mathbf{R}_u=\mathbf{C}_u-\mathbf{C}_u\mathbf{F}_u^*(\mathbf{F}_u\mathbf{C}_u\mathbf{F}_u^*+\mathbf{C}_p)^{-1}\mathbf{F}_u\mathbf{C}_u. \tag{47}$$

Setting $\mathbf{C}_p$ and $\mathbf{C}_u$ to scaled identity matrices, setting σ_p^2 to the mean squared residual error (i.e., the data-mode misfit from the adjoint-based iterative inversion after 50 iterations), and setting σ_u^2 so that β was 25, we calculated $\mathbf{R}_u$ using Eq. (46). Using the residual errors to set the prior covariance would not be justifiable with real data (this variance should represent *a priori* knowledge of the measurement system accuracy), but seemed a reasonable way in this illustrative simulated example to set a representative level for these variances.

Figure 13 shows the true sound speed, the adjoint estimate for the sound speed, and bounding curves three standard deviations above and below the adjoint estimate. The standard deviations are the diagonal elements of $\mathbf{R}_u$. The eigenvectors of $\mathbf{R}_u$ are the principal components of the errors, and can be plotted to show the spatial correlation of the errors. Because $\mathbf{C}_u$ is diagonal, its addition only serves to

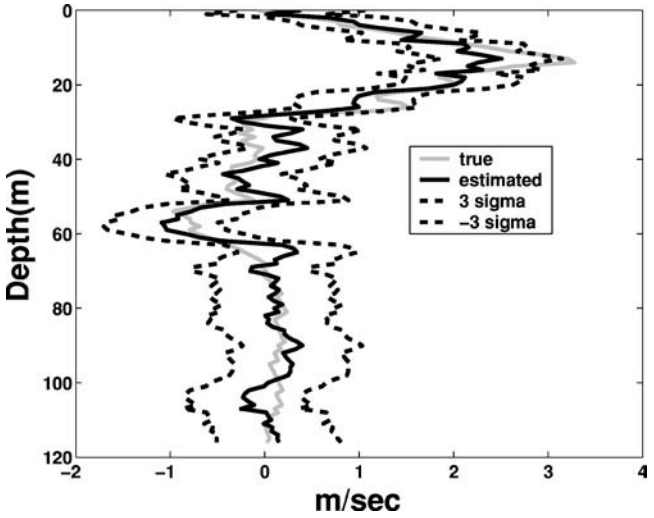


FIG. 13. Sound speed estimates with error bars (relative to baseline sound speed profile).

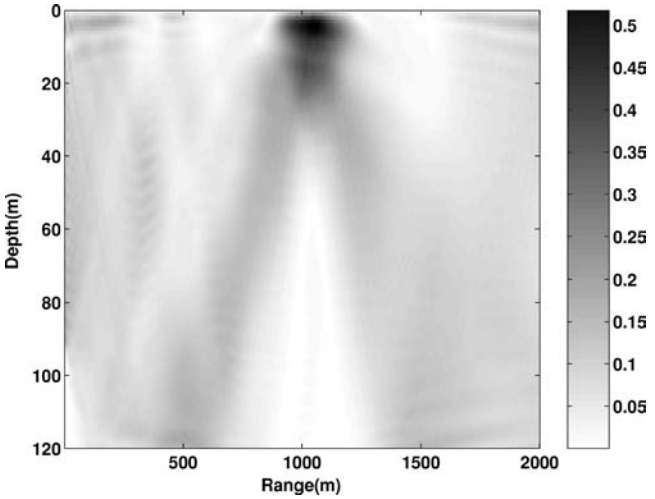


FIG. 15. Image of 2D feature at 1000 m after a single iteration of the adjoint process. Image scale shows magnitude of the correction to be applied to the sound speed field.

modify the eigenvalues of $\mathbf{R}_u$ and not its eigenvectors, causing the eigenvalues of $\mathbf{R}_u$ to increase or decrease according to changes in σ_u^2.

E. Adjoint inversions (range-dependent case)

In Sec. III B, we inverted for range-independent sound speed perturbations. In this section, we invert for range-dependent sound speed perturbations. To construct a range-dependent test case, we treat the time sequence of INTIMATE 96 sound speed profiles as a sequence of profiles spanning a range interval. Figure 14 shows the range-dependent sound speed perturbation we will be imaging with our adjoint model (this is a perturbation from the mean profile). This 2D sound speed perturbation was generated in two steps. First, the sequence of INTIMATE 96 profiles was interpolated to cover 2 km in range with a 1-m spacing between profiles. Second, we isolated a short range interval where a large deviation from the mean profile occurred, by applying a Gaussian weighting centered at this interval. This interval covers profiles 16–18 in Fig. 11. Working with an isolated 2D feature enabled us to translate this structure to different ranges (500, 1000 and 1500 m), so that we could compare our ability to image it at these different ranges.

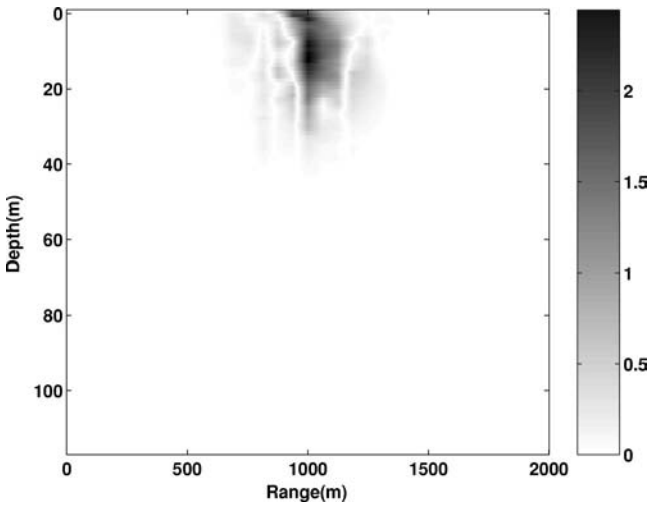


FIG. 14. Range-dependent sound speed perturbation being imaged by adjoint process. This figure shows the range-dependent feature being imaged at 1000 m. Image scale shows magnitude of perturbation in sound speed measured in m/s.

The parabolic equation model described in Sec. II was used to generate synthetic acoustic data on the receive array used in previous sections, for sources at depths from 5 to 90 m in 5-m steps at 100–400 Hz in 5-Hz steps. In contrast to the results in Sec. III B, where the calculated sound speed corrections were averaged over range (since range-independent sound speed fields were being inverted for), each source depth and frequency here produced a 2D map of sound speed corrections versus range and depth. Several combinations of source depths and source frequencies were used to calculate inversions for the range-dependent sound speed perturbation shown in Fig. 14. Using only a single source depth resolved the location of the range-dependent feature, but did not resolve the shape very well, producing only several highlights within the shape. The positions of these highlights changed with source depth. Using multiple frequencies did not improve the results as much as using multiple source depths, probably because it is more important to adequately sample the feature with acoustic paths than to use a wider bandwidth. Changing the frequency of the ranging signal does not appreciably change the acoustic paths visiting the 2D feature being imaged.

Figure 15 shows the summation of maps from all source depths (5–90 m in 5-m increments) and all frequencies (100–400 Hz in 5-Hz increments). Only a single iteration of the process applied in Sec. III B was used. Figures 14 and 15 show the adjoint model has imaged the 2D feature at 1000 m. Further iterations would better resolve the shape and produce an amplitude that better matches that shown in Fig. 14. This feature was similarly imaged when translated to 500 and 1500 m (although these results are not shown).

This shows that an adjoint model can produce a 2D map of large dimension using a single pass of the adjoint model. However, there is no guarantee that the map of corrections

produced by the adjoint model is unique. With too few measurements (e.g., too few source depths), it is conceivable that many 2D maps of corrections could be found to reproduce the same data-model misfit at the receive array. If the inverse problem has multiple solutions, the adjoint process provides no indication that a solution it has found is not unique. Analysis of the matrix used to linearize the problem [e.g., Eq. (19)] and its null space could be used to alert us to this ambiguity.

These results also rely upon the range-dependent sound speed perturbation being small enough that our initial guess (the mean profile, as defined in previous sections) is in some neighborhood of the objective function minimum. An initial guess too far away from the true profile could result in the inversion process getting trapped in a spurious local minimum.

IV. CONCLUSIONS

As we have shown above, an adjoint model propagates the data-model misfit (which is often identified as the adjoint field) back through the medium and calculates the perturbations to the forward model inputs needed to correct for this misfit. The adjoint model generates a sensitivity map of the model parameters $\mathbf{u}_n$ to the observations $\mathbf{m}_N$, providing a direction for steepest descent in inverse problems. The alternative is to repeatedly run the forward model to test how each possible perturbation in the model parameters $\mathbf{u}_n$ influences the observables $\mathbf{m}_N$. The adjoint model can thus provide a more economical way of exploring the search space (of model parameters $\mathbf{u}_n$) in aid of solving inverse problems. The computational savings of using an adjoint model increases with the number of unknowns in the problem being addressed, since each unknown in the problem presents a potentially distinct perturbation that must be assessed by the forward model. As a result, inversion processes using an adjoint model can accommodate large-dimensional problems, without demanding lower-dimensional representations, often required by alternative inverse methods in order to reduce the number of forward modeling runs needed to explore the search space.

To show how an adjoint model arises in the familiar setting of solving a linear system, we have shown how first-order perturbations to the PE marching solution can be formulated as a linear system and solved using a steepest descent iteration, where the adjoint of the linear system is used to calculate the gradient. The steepest-descent iteration provides an alternative to calculating the inverse of a typically very large matrix. We have also presented an alternative formulation of the adjoint, showing how the adjoint model for the marching PE solution can be derived using Lagrange multipliers, which are the ''adjoint field.'' The adjoint model back-propagates its field from the measurement locations back to the medium perturbations that are causing data-model misfit. Because the ''adjoint field'' provides a mechanism for back-propagating information in this way, there is no need to construct the potentially very large linear system matrix, needed in the former derivation. Instead, we use two models, the original PE forward model and its adjoint, to propagate their respective fields in their respective directions. A fixed-point iteration results from this formulation, which we have shown is equivalent to a steepest-descent iteration based on a gradient, calculated from the ''adjoint field'' back-propagated by the adjoint model.

As a demonstration of these techniques, we have applied our fixed-point iteration to the series of sound speed profiles measured during the INTIMATE 96 experiment. We use acoustic ''data'' calculated by our PE model with the measured sound speed profiles provided as inputs. We invert the (synthetic) acoustic ''data'' for the sound speed profiles, to demonstrate our technique. We have shown successful inversions for both range-independent and range-dependent sound speed features. For the range-independent case we have also shown how to provide estimates of the accuracy of our results.

The adjoint-assisted inversion presented in this paper can be viewed as a wave-theoretic approach to tomography, and thus may provide an alternative to ray-based tomography and its shortcomings.[30] However, we recognize that relying upon pressure as the observable quantity is problematic, and hope in the future to extend the adjoint formulation to more robust observables.

Lest we leave the impression that adjoint modeling is a panacea for all inverse problems, we describe some difficulties that must be overcome in adjoint modeling. An adjoint model only works perfectly when the analogous forward model is linear. This has not inhibited the use of adjoint models in other fields, since mildly nonlinear problems can typically be linearized and more than slightly nonlinear problems can be attacked with iterative formulations[14,31] in which an adjoint model remains a valuable component. Nevertheless, some problems will not lend themselves to adjoint modeling, because of their inherent nonlinearity. The adjoint method finds a local minimum, with no guarantee of it being globally optimal in problems whose objective function is not convex. Like other local methods, the adjoint iterative process may require a more sophisticated optimization strategy (e.g., conjugate gradients) if the simple steepest descent presented in the paper converges too slowly (the fixed point iteration may not converge at all). Finally, an adjoint model poses the burden of requiring additional implementations of both a tangent linear and an adjoint model, over and above the implementation of the forward model. Efforts are underway to automate the construction of such models from the original forward model code.[32]

ACKNOWLEDGMENTS

The paper was much improved by the suggestions of the anonymous reviewers. This work was supported by the Office of Naval Research, Grant No. N00014-01-1-0768.

[1] A. Tarantola, ''Inversion of seismic reflection data in the acoustic approximation,'' Geophysics **49(8)**, 1259–1266 (1984).
[2] O. Dorn, E. L. Miller, and C. M. Rappaport, ''A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets,'' Inverse Probl. **16**, 1119–1156 (2000).
[3] W. Munk, P. Worcester, and C. Wunsch, *Ocean Acoustic Tomography* (Cambridge U.P., New York, 1995), pp. 318–319.
[4] C. Wunsch, *The Ocean Circulation Inverse Problem* (Cambridge U.P., New York, 1996), pp. 362–391.

[5] A. F. Bennett, *Inverse Methods in Physical Oceanography* (Cambridge U.P., New York, 1992), pp. 112–135.
[6] S. J. Norton, ''Iterative inverse scattering algorithms: Methods of computing Fréchet derivatives,'' J. Acoust. Soc. Am. **106**, 2653–2660 (1999).
[7] M. Asch, J.-C. L. Gac, and P. Helluy, ''An adjoint method for geoacoustic inversions,'' in *Proceedings of the 2nd Conference on Inverse Problems, Control and Shape Optimization* (Carthage, Tunisia, 2002).
[8] A. Thode, ''An adjoint/normal-mode approach for computing multiple-order derivatives of pressure with respect to environmental parameters and frequency: range-independent waveguide,'' J. Acoust. Soc. Am. , (2003) (submitted).
[9] M. D. Collins and W. A. Kuperman, ''Focalization: Environmental focusing and source localization,'' J. Acoust. Soc. Am. **90**, 1410–1422 (1991).
[10] P. Gerstoft, ''Inversion of seismoacoustic data using genetic algorithms and *a posteriori* probability distributions,'' J. Acoust. Soc. Am. **95**, 770–782 (1994).
[11] A. Tolstoy, N. Chapman, and G. Brooke, ''Workshop '97: Benchmarking for geoacoustic inversion in shallow water,'' J. Comput. Acoust. **6**(1), 1–28 (1998).
[12] A. Tolstoy, ''Linearization of the matched field processing approach to acoustic tomography,'' J. Acoust. Soc. Am. **91**, 781–787 (1992).
[13] V. V. Goncharov and A. G. Voronovich, ''An experiment on matched-field acoustic tomography with continuous wave signals in the Norway Sea,'' J. Acoust. Soc. Am. **93**, 1873–1881 (1993).
[14] D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969).
[15] P. R. McGillivray and D. W. Oldenburg, ''Methods for calculating Fréchet derivatives and sensitivities for the non-linear inverse problem: a comparative study,'' Geophys. Prospect. **38**, 499–524 (1990).
[16] D. E. Kirk, *Optimal Control Theory* (Prentice–Hall, Englewood Cliffs, NJ, 1970), pp. 184–240.
[17] W. L. Brogan, *Modern Control Theory*, 3rd ed. (Prentice–Hall, Englewood Cliffs, NJ, 1991).
[18] I. M. Gelfand and S. V. Fomin, *Calculus of Variations* (Prentice–Hall, Englewood Cliffs, NJ, 1963), pp. 218–225.
[19] A. F. Bennett, *Inverse Modeling of the Ocean and Atmosphere* (Cambridge U.P., New York, 2002), pp. 18–29.
[20] A. J. Devaney, ''A filtered backpropagation algorithm for diffraction tomography,'' Ultrason. Imaging **4**, 336–350 (1982).
[21] E. Coelho, ''Mesoscale-small scale oceanic variability effects on underwater acoustic signal propagation,'' in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 49–54.
[22] H. A. Deferrari, N. J. Williams, and H. B. Nguyen, ''Variability, coherence and predictability of shallow water acoustic propagation in the straits of Florida,'' in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 245–254.
[23] A. R. Robinson, P. Abbott, P. F. J. Lermusiaux, and L. Dillman, ''Transfer of uncertainties through physical-acoustical-sonar end-to-end systems: a conceptual basis,'' in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 409–416.
[24] A. Warn-Varnas, S. Chin-Bing, D. King, J. Hawkins, K. Lamb, and M. Teixeira, ''Yellow Sea internal solitary wave variability,'' in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 409–416.
[25] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP, Woodbury, NY, 1994), pp. 366–375.
[26] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations* (Chapman and Hall, New York, 1989), pp. 323–350.
[27] Y. Stéphan, X. Démoulin, T. Folégot, S. M. Jesus, M. B. Porter, and E. Coelho, ''Acoustical effects of internal tides on shallow water propagation: an overview of the INTIMATE96 experiment,'' in *Experimental Acoustic Inversion Methods for Exploration of the Shallow Water Environment*, edited by A. Caiti, J. P. Hernand, S. M. Jesus, and M. B. Porter (Kluwer Academic, Dordrecht, The Netherlands, 2000), pp. 19–38.
[28] The singular values were within 1% of each other out to 110 of 117 of the eigenvectors. Similarly, the eigenvectors were correlated to within 1% out to 80 of 117 eigenvectors, and to within 5% out to 110 of 117 eigenvectors. The singular values for eigenvectors 80 and 110 were 60 and 80 dB down from the first eigenvector, respectively.
[29] A. Tarantola, *Inverse Problem Theory* (Elsevier Science, New York, 1987), pp. 187–294.
[30] G. A. Athanassoulis and E. K. Skarsoulis, ''Arrival-time perturbations of broadband tomographic signals due to sound-speed disturbances. A wave-theoretic approach,'' J. Acoust. Soc. Am. **97**, 3575–3588 (1995).
[31] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, San Diego, CA, 1981).
[32] R. Giering and T. Kaminsky, ''Recipes for adjoint code construction,'' ACM Trans. Math. Softw. **24**(4), 437–474 (1998).

Matched-field depth estimation for active sonar

Granger Hickman and Jeffrey L. Krolik[a)]
Duke University, Department of Electrical and Computer Engineering, Box 90291, Durham, North Carolina 27708-0291

(Received 11 August 2000; revised 21 October 2003; accepted 23 October 2003)

This work concerns the problem of estimating the depth of a submerged scatterer in a shallow-water ocean by using an active sonar and a horizontal receiver array. As in passive matched-field processing (MFP) techniques, numerical modeling of multipath propagation is used to facilitate localization. However, unlike passive MFP methods where estimation of source range is critically dependent on relative modal phase modeling, in active sonar source range is approximately known from travel-time measurements. Thus the proposed matched-field depth estimation (MFDE) method does not require knowledge of the complex relative multipath amplitudes which also depend on the unknown scatterer characteristics. Depth localization is achieved by modeling depth-dependent relative delays and elevation angle spreads between multipaths. A maximum likelihood depth estimate is derived under the assumption that returns from a sequence of pings are uncorrelated and the scatterer is at constant depth. The Cramér–Rao lower bound on depth estimation mean-square-error is computed and compared with Monte Carlo simulation results for a typical range-dependent, shallow-water Mediterranean environment. Depth estimation performance to within 10% of the water column depth is predicted at signal-to-noise ratios of greater than 10 dB. Real data results are reported for depth estimation of an echo repeater to within 10-m accuracy in this same shallow water environment.
[DOI: 10.1121/1.1634277]

PACS numbers: 43.30.Vh, 43.30.Wi, 43.30.Gv [DLB] Pages: 620–629

I. INTRODUCTION

Matched-field processing (MFP) is a technique for estimating parameters such as the range, depth, and bearing of a source in an ocean acoustic waveguide. MFP estimates the target's location by comparing the received data with predictions generated by a propagation model over a set of hypothesized target locations. By utilizing a numerical propagation model, it is possible to exploit acoustic multipath complexities brought about by the ocean environment rather than trying to mitigate them, as is often done when using classical plane wave beamforming. First formulated for passive sonar by Bucker, an extensive literature on MFP techniques has developed since Fizell and Wales first demonstrated the feasibility of MFP with field data.[1,2] Since 1985, other experiments have appeared in the literature that also support the feasibility of matched-field processing approaches, given sufficiently accurate environmental modeling and an adequate signal-to-noise ratio. These results have been surveyed by Baggeroer *et al.*[3]

Although much work has been done in MFP, the majority of this effort has been applied to passive sonar. Application of MFP to active sonar is complicated by the presence of the scatterer, which can modify the amplitude and phase of each multipath return returned to the receiver. A further complication in using active sonar with traditional MFP is that if the sonar pings are time-limited and well-separated (to avoid range ambiguities), then it is difficult to apply adaptive array processing methods which require adequate sample support in order to estimate the cross-spectral data matrix at each ping. Despite the emphasis on passive sonar in MFP, there has been work done on understanding the effect of the ocean waveguide on active sonar signals and tracking targets with active sonar systems.[4–6] Yang and Yates have presented a broadband approach to MFP localization, built upon a normal mode scattering formulation, that can localize a target that disturbs the relative magnitudes (but not the relative phases) between the multipaths that make up a ping return: as Yang and Yates note, this assumption may be realistic for certain applications.[7,8] The approach proposed in this paper assumes that the target's scattering function is random but constant over the bandwidth of the insonifying pulse. In other words, the overall amplitude and phase of each multipath return that makes up a single ping is randomly changed by the scatterer. Moreover, the complex amplitudes of the multipath returns across a series of pings are assumed to be uncorrelated, zero-mean Gaussian random variables with known variance. With this model, the main information available regarding scatterer location is the group delays of multipath returns. The assumption that the scattering function is constant over the bandwidth of the transmitted waveform necessitates the use of relatively narrow-band sonar waveforms, resulting in unresolved multipath propagation delays. It can be seen that conventional MFP, the MFP approach of Yang and Yates, and the MFP approach proposed in this paper represent three levels of restriction placed on the scatterer. Conventional narrow-band MFP assumes that the scatterer does not change the relative amplitudes or phases of the multipaths or modes, the Yang and Yates approach can localize the target even if the relative amplitudes

[a)]Author to whom correspondence should be addressed. Electronic mail: jk@ee.duke.edu

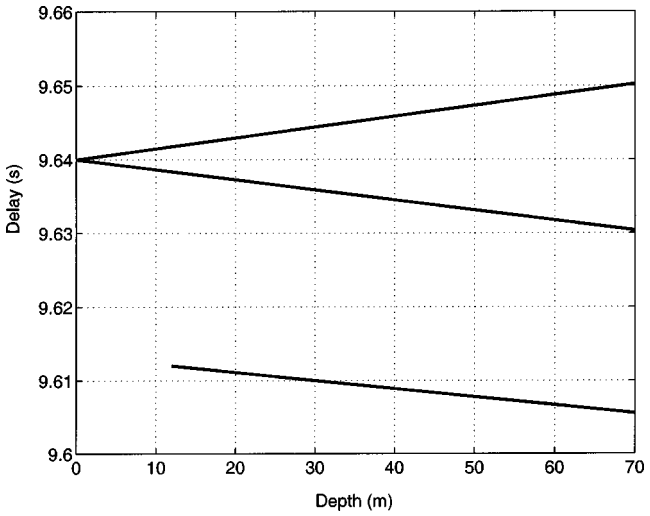


FIG. 1. Depth versus delay. Each vertical slice shows the number of rays arriving at that depth while at 14.5 km from the source. (For instance, two rays arrive at 10 m in depth, each with a delay of about 9.64 s.) Note that the largest delay spread (approximately 0.05 s) is smaller than the duration (0.5 s) of the transmitted signal.

are changed by the target, and the maximum likelihood approach considered in this paper assumes only knowledge of the statistics of the scattering function.

In active systems the scatterer range is approximately known from round trip group delay measurement, and bearing can be estimated by mapping field directionality using a horizontal or cylindrical array. Scatterer depth, however, is not easily discernible with active sonar. Depth estimation is achieved in this paper by exploiting the depth dependence of the relative multipath delays. For example, in Fig. 1, the depth dependence of delay spread is illustrated for a Mediterranean environment (summarized in Fig. 2) and target-receiver scenario in Sec. IV. It can be seen from this figure that the number of multipaths, as well as the relative delays between them, is a function of target depth (and range, which is assumed to be known here). In essence, a depth-dependent linear signal model is used to fit the complex frequency-wavenumber return of the scatterer. Because of the limited source bandwidth, single-ping depth accuracy is low, but when incoherently accumulated across an uncorrelated sequence of pings, it is shown (for a typical Mediterranean environment) that maximum likelihood (ML) MFDE can achieve accuracy to within 10% of the water column depth at signal-to-noise ratios above 10 dB. The reason for the improvement in estimate as the number of pings increases is the asymptotic convergence property of ML estimators.[9] In comparison with passive MFP, the availability of a bulk group delay estimate in active sonar avoids many of the serious localization ambiguities inherent in passive methods which estimate target range from relative multipath phase differences.

The remainder of this paper is organized as follows. A model for the target return, reverberation, and noise is proposed in Sec. II. In Sec. III this model is used to derive a time-evolving maximum likelihood depth estimator. Finally, Sec. IV examines the application of the depth estimator to simulations and to real transponder data collected during a Mediterranean experiment.

II. MODELING MULTIPATH ACTIVE SONAR RETURNS

Consider a uniformly spaced horizontal linear hydrophone array towed at a constant depth which receives active sonar pings scattered from a target. The sonar source may be co-located with the receiver, or alternately may be at a different location and depth, and they need not necessarily maintain a fixed relationship with respect to each other. For the purposes of this paper, the target is a scatterer assumed to maintain a constant depth over the collected sequence of pings.

Using a ray-based propagation model, it can be seen that acoustic signals propagate from the source to the target along multipaths that are determined by the depths of the source, target, and receiver, as well as by sound speed profile and ocean bathymetry. The term ''multipath'' refers here to the two-way ray combination from the source, to the target at depth z, then to the receiver. A simple example of one-way multipath propagation is illustrated by the ray trajectories

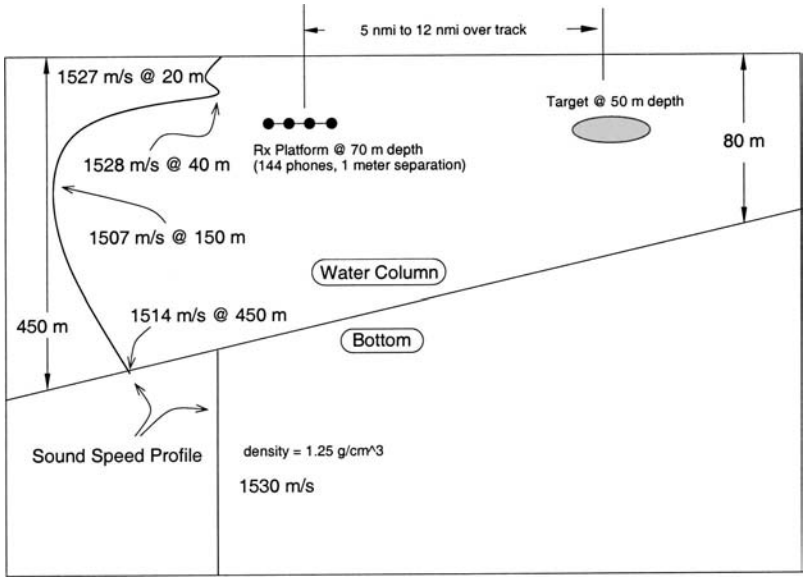


FIG. 2. The scenario and environment of the Mediterranean experiment. The left side of the figure summarizes the salient features of the sound speed profile, while the right side shows the target-receiver geometry relative to the bathymetry. Note that the water column is shallower near the target.

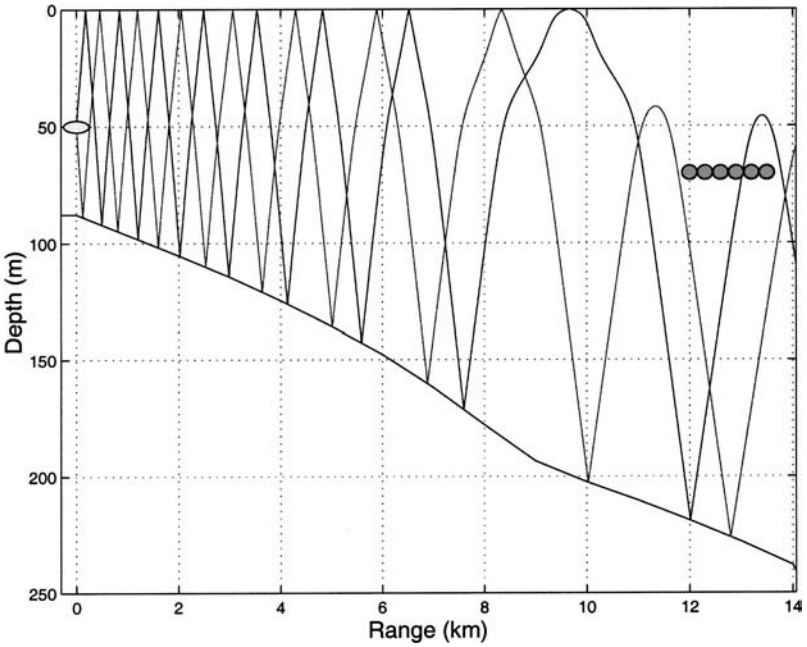


FIG. 3. Illustration of multipath ray propagation through an ocean with the Mediterranean experiment's environmental parameters, with rays propagating from the target (left) to the receiver (right). Only two rays are shown for visual clarity.

shown in Fig. 3. Each incoming ray at the target can be back-scattered to the receiver array along another (generally different) set of rays between the target and the receiver. Note that the set of source-target rays will be identical to the set of target-receiver rays if the source and receiver are co-located.

The signal transmitted along a single ray path to a single element of the receiver is generally time-delayed, Doppler-shifted, and spread due to target scattering effects. The principal ray propagation quantities that will be modeled in this paper are the delays for each of the multipaths, and the launch and arrival angles at the multipath termini. The reflection coefficient, as a function of angle, is assumed to be frequency-independent at the ocean boundaries which is appropriate as long as the sound speed and density properties of the water and ocean bottom are frequency-independent.[10] Since the transmitted signals considered in this paper have bandwidths of 10 Hz or less, this assumption is reasonable for the purposes considered here. Let $w(t)$ denote the original transmitted signal (where the variable t represents time). Further, let $\tau_{ij}(z)$ represent the propagation delay (for a target at depth z) incurred by traveling along the multipath consisting of the ith source-target ray ($i \in \{1,...,I\}$) and the jth target-receiver ray ($j \in \{1,...,J\}$), where the target-receiver ray joins the target and first receiver element of the array. If, for the signal traveling along this multipath, $f_{ij}(z)$ is defined as the total complex amplitude change due to the boundary interactions and ray propagation, then the contribution seen at the first receiver element is given by $w(t) \star (f_{ij}(z)\delta(t - \tau_{ij}(z))) = f_{ij}(z)w(t-\tau_{ij}(z))$, where the operator "$\star$" represents temporal convolution. To get the signal seen by the pth receiver element ($p \in \{1,...,P\}$), the delay must be adjusted by $\cos\gamma \cos\phi_j(z)(d/c)(p-1)$, where γ is the azimuthal angle measured from array end-fire to the direction of the target, $\phi_j(z)$ is the elevation arrival angle of the jth target-receiver ray, d is the receiver element separation, and c is the sound speed at the depth of the receiver. These angles are diagrammed in the top and bottom panels, respectively, of Fig. 4. Another convolution, similar to the one above, leads to the expression

$$\begin{aligned} g_{ij}(t,p,z) &= f_{ij}(z)w(t-\tau_{ij}(z)) \star \delta(t - \cos\gamma \cos\phi_j(z) \\ &\quad \times (d/c)(p-1)) \\ &= f_{ij}(z)w\Big(t - \tau_{ij}(z) - \cos\gamma \\ &\quad \times \cos\phi_j(z)\left(\frac{d}{c}\right)(p-1)\Big) \end{aligned} \tag{1}$$

for the contribution at the pth receiver element for the multipath consisting of the ith source-target ray and the jth target-receiver ray. Notice that this treatment of $\phi_j(z)$ assumes that each ray arriving at the receiver array is like an incident plane wave. This assumption should hold good for an array that is short enough, and it reduces the amount of ray tracing necessary to approximate the field seen at the receiver array. The quantities $f_{ij}(z)$, $\tau_{ij}(z)$, and $\phi_j(z)$ are estimated numerically by ray tracing.

Equation (1) neglects the effects of the scattering target on the received signal. The effect of the target on the portion of the signal traveling along the multipath mentioned in the previous paragraph is approximated here by the action of a linear, time-invariant filter, $b_{ij}(t,z)$, which is dependent on the particular multipath and the depth of the target. If the target size is small compared to the length of the signal (i.e., if $l \ll cT/2$, where l is the physical length of the target, c is sound speed, and T is the duration of the signal), then the signal scattered from a source-target path has approximately the same time duration as the incident signal arriving from a source-target path. This is consistent with the behavior of a relatively narrow-band source with $T \gg 2l/c$, so that $b_{ij}(t,z) \approx s_{ij}\delta(t)$, where the scalar, complex scattering coefficient, s_{ij}, is dependent on the multipath indexed by i and j.

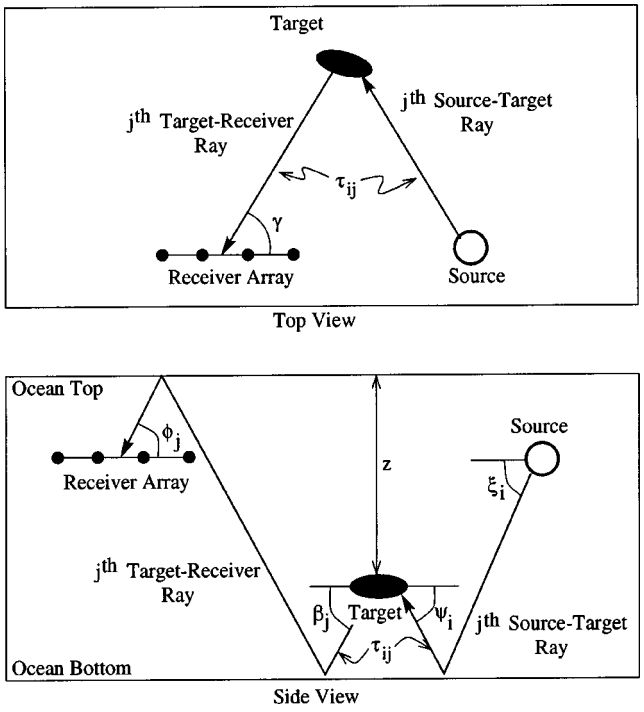


FIG. 4. Diagrams illustrating various ray parameters. Top view: The bearing angle γ is measured from the end-fire direction of the array to the line joining the receiver array center and target. Side view: The elevation arrival angle of the jth ray, ϕ_j, is measured from the end-fire direction of the array to the line tangent to the ray as it crosses the array center. The elevation source departure angle, ξ_i, elevation target arrival angle, ψ_i, and the elevation target departure angle, β_j, are measured similarly. The group delay, τ_{ij}, which is indicated in both plots, is the group delay seen as the signal traverses the multipath comprised of the ith source-target ray and the jth target-receiver ray. If there are I source-target rays and J target-receiver rays, then there are, in general, $I\times J$ different group delays and multipaths. For clarity, only one such multipath is indicated in the side view.

The additional delay introduced by the scattering process is assumed to be negligible. An important point here is that s_{ij} is a nonconstant function of the particular multipath (indexed by i and j). Thus, accounting for propagation effects and scattering, the approximate signal contribution at the pth receiver element is, after convolving $b_{ij}(t,z)$ with the right-hand side of Eq. (1),

$$h_{ij}(t,p,z)=g_{ij}(t,p,z)\star b_{ij}(t,z)$$
$$=s_{ij}f_{ij}(z)w\Big(t-\tau_{ij}(z)$$
$$-\cos\gamma\cos\phi_j(z)\frac{d}{c}(p-1)\Big). \tag{2}$$

The scattering coefficients s_{ij}, which are parameters of the target, are unknown because the precise nature of the target is unknown. The chief aim of Sec. III is to derive an estimator of the target depth, z, in the absence of specific knowledge of the values of the target's scattering coefficients.

Next, the Doppler, which is induced by motion through the acoustic medium, is taken into account. The receiver, target, and source speeds, relative to the acoustic medium, are necessary for calculation of the Doppler on the transmitted signal. Define the horizontal speed at which the receiver moves in the direction of the target as v_{rt} (and vice versa as v_{tr}), and the speed at which the source moves in the direction of the target as v_{st} (and vice versa as v_{ts}). Referring to the bottom half of Fig. 4, denote $\phi_j(z)$ and $\beta_j(z)$ as the elevation angles of arrival and departure, respectively, for the jth target-receiver ray at the pth receiver element. Similarly, define $\psi_i(z)$ and $\xi_i(z)$ as the elevation angles of arrival and departure, respectively, for the ith source-target ray. Finally, denote the speed of sound at the receiver, target, and source depths as c_r, c_t, and c_s, respectively. Taking all of the above quantities, for a signal traveling along the ith source-target ray and the jth target-receiver ray to a receiver element,

$$\alpha_{ij}(z)=1+\frac{v_{rt}}{c_r}\cos\phi_j(z)+\frac{v_{tr}}{c_t}\cos\beta_j(z)+\frac{v_{ts}}{c_t}\cos\psi_i(z)$$
$$+\frac{v_{st}}{c_s}\cos\xi_i(z) \tag{3}$$

is the total Doppler factor.[11] Putting the Doppler term from Eq. (3) into Eq. (2) yields the result

$$u_{ij}(t,p,z)=h_{ij}(\alpha_{ij}(z)t,p,z). \tag{4}$$

Note that the various angles used in Eq. (4) must be obtained numerically via ray tracing.

Equations (1)–(4) summarize the continuous time-domain model for the contribution of a single multipath at a single receiver. The discrete signal is sampled at time indices, n, with sampling interval T_0. Further, the signal at receiver p will be corrupted with noise $\nu(n,p)$, assumed here to be white, zero mean, and complex with Gaussian distribution. The discrete time-domain model for the total signal $y(n,p,z)$ seen at receiver element p due to a target at depth z is then

$$y(n,p,z)=\sum_{i,j}u_{i,j}(nT_0,p,z)+\nu(n,p)$$
$$=\sum_{i,j}s_{ij}q_{i,j}(n,p,z)+\nu(n,p), \tag{5}$$

where

$$q_{i,j}(n,p,z)=f_{ij}(z)w(\alpha_{ij}(z)nT_0-\tau_{ij}(z)$$
$$-\cos\gamma\cos\phi_j(z)(d/c)(p-1)).$$

In order to suppress interference in both space and time, the data modeled in (5) is typically windowed and discrete Fourier transformed, after which only the target beams and frequency band are retained. Letting $v(n,p)$ denote the two-dimensional window, the resulting frequency-wavenumber description is given by $U(\omega,\kappa,z)=(V\star Y)(\omega,\kappa,z)$, where $\omega\in\{1,...,\Omega\}$ and $\kappa\in\{1,...,K\}$ are the temporal and spatial digital frequencies, respectively, and the DFT of $v(n,p,z)$ and $y(n,p,z)$ is $V(\omega,\kappa,z)$ and $Y(\omega,\kappa,z)$, respectively. The "$\star$" operation here indicates convolution in both the frequency index ω and the wavenumber index κ. The complex frequency-wavenumber spectrum of the data can then be written as

$$U(\omega,\kappa,z)=\sum_{i,j}s_{ij}(V\star Q_{ij})(\omega,\kappa,z)+\eta(\omega,\kappa), \tag{6}$$

where $Q_{ij}(\omega,\kappa,z)$ denotes the DFT of $q_{ij}(n,p,z)$, and $\eta(\omega,\kappa)$ is the DFT of $\nu(n,p)$. Since the signal energy is assumed to be band-limited in both frequency and wavenumber, the information represented by U can be approximated by an $F \times W$ subset of frequency-wavenumber bins that cover the energy-extent of the signal, where F is the size of the frequency index subset and W is the size of the wavenumber index subset. In other words, if $\{\omega_1,\omega_2,\ldots,\omega_F\}$ is a set of F frequency indices that span the frequency bandwidth of the signal, and $\{\kappa_1,\kappa_2,\ldots,\kappa_W\}$ is a set of W wavenumber indices that span the wavenumber bandwidth of the signal, then the signal information is largely contained in the $F\times W$ matrix $\mathbf{U}(z)$, where a given coefficient of $\mathbf{U}(z)$ can be characterized as $(\mathbf{U}(z))_{fw}=\Sigma_{i,j}s_{ij}(V\star Q_{ij})(\omega_f,\kappa_w,z)+\eta(\omega_f,\kappa_w)$, $f \in\{1,...,F\}$, and $w\in\{1,...,W\}$. For convenience, define matrices $\mathbf{H}_{ij}(z)$ and $\mathbf{N}$ such that

$$(\mathbf{H}_{ij}(z))_{fw}=(V\star Q_{ij})(\omega_f,\kappa_w,z) \tag{7}$$

and

$$(\mathbf{N})_{fw}=\eta(\omega_f,\kappa_w) \tag{8}$$

for all $f\in\{1,...,F\}$ and $w\in\{1,...,W\}$. It follows that for returns from a single transmitted ping, the relevant portion of the frequency-wavenumber surface, $\mathbf{U}(z)$, can summarized as a $FW\times 1$ vector:

$$\begin{aligned}\mathbf{x}(z)&=\mathrm{vec}(\mathbf{U}(z))=\sum_{i,j}\mathrm{vec}(\mathbf{H}_{ij}(z))s_{ij}+\mathrm{vec}(\mathbf{N})\\&=\mathbf{H}(z)\mathbf{s}+\mathbf{n},\end{aligned} \tag{9}$$

where the operator vec(•) stacks the columns of a matrix to form a column vector, $\mathbf{H}(z)$ is a $(FW)\times(IJ)$ matrix whose columns (each representing a single multipath return) are given by $\mathrm{vec}(\mathbf{H}_{ij}(z))$, and $\mathbf{s}$ is a complex $(IJ)\times 1$ vector whose elements are given by s_{ij}. The scattering vector, $\mathbf{s}$, describes the target's effect on the received data, $\mathbf{x}$, and is treated as a zero-mean Gaussian random vector.

Equations (5)–(9) neglect the reverberation caused by the active sonar, but, as illustrated in Fig. 5, reverberation may be included by assuming that the transmitted signal is also scattered to the receiver by L patches on the ocean floor.[12] The lth reverberation patch, which lies at a depth of z_l and has the same time-of-flight as the target, scatters the transmitted signal to the receiver. This patch has a replica matrix $\mathbf{H}_l(z_l)$ associated with it (with a derivation analogous to that of $\mathbf{H}(z)$ from Eq. (9)), and each column of this replica matrix represents the frequency-wavenumber information associated with a particular two-way multipath involving the source, the receiver, and the lth patch. Similarly, there is a scattering vector $\mathbf{g}_l$ associated with the lth reverberation patch, which arises in the same manner as the scattering vector $\mathbf{s}$ does for the target. The length of $\mathbf{g}_l$ equals the number of columns $\mathbf{H}_l(z_l)$, and each coefficient of the scattering vector $\mathbf{g}_l$ represents the complex amplitude imposed on the corresponding multipath's contribution at the receiver. Thus, including reverberation terms yields the data model

$$\mathbf{x}=\mathbf{H}(z)\mathbf{s}+\sum_{l=1}^{L}\mathbf{H}_l(z_l)\mathbf{g}_l+\mathbf{n}. \tag{10}$$

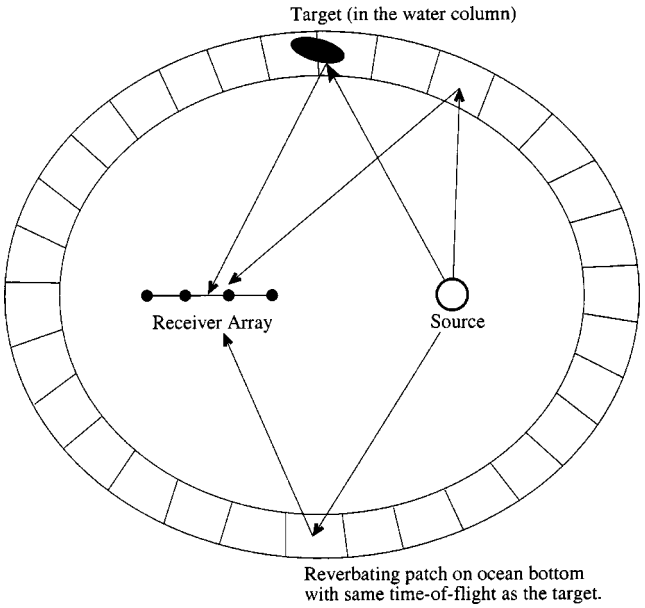


FIG. 5. Diagram of the reverberation ellipse, divided into L reverberation patches. The reverberation ellipse consists of reverberating locations on the ocean bottom that have the same time-of-flight (from the source to the receiver) as the target.

To summarize, the data model, given by Eq. (10), consists of known parameters, unknown parameters, and random vectors. The parameters necessary for calculating the matrix $\mathbf{H}$ (summarized by Eqs. (7) and (9)) are the group delays and ray angles [Eqs. (2)–(4)], which can be computed with a ray trace model with inputs of ocean environment and bathymetry, as well as the positions and velocities of the source, receiver and target. Except for the target's depth, z, all of the parameters of $\mathbf{H}(z)$ are assumed to be known. The target depth, z, is the unknown nonrandom parameter of interest. The additive noise vector, $\mathbf{n}$, from Eq. (10), is taken to be a zero-mean Gaussian random vector. Likewise, the scattering vectors, $\mathbf{s}$ and $\mathbf{g}_l$, are modeled as zero-mean Gaussian random vectors. Modeling of the relationship between the returns from a sequence of pings, $\mathbf{x}_k$, where $k\in\{1,...,K\}$, is done in Sec. III along with the discussion of a maximum likelihood estimate of target depth.

III. MAXIMUM LIKELIHOOD DEPTH ESTIMATION WITH MULTIPLE PINGS

In this section a maximum likelihood estimator is discussed that uses multiple active sonar pings and assumes a statistical distribution for the scattering characteristics of the target. Consider a target which moves in the ocean at a constant depth. The sonar source and receiver will also be assumed to move at constant depths. Using Eq. (10), let $\mathbf{x}_k=\mathbf{H}_k(z)\mathbf{s}_k+\Sigma_{l=1}^{L}\mathbf{H}_{lk}(z_{lk})\mathbf{g}_{lk}+\mathbf{n}_k$ denote the complex frequency-wavenumber data vector from the kth ping (where $k\in\{1,...,K\}$ represents the ping index). A conventional choice (i.e., a choice appropriate for conventional MFP) for the scattering vector, $\mathbf{s}_k$, would be a constant vector with all entries set to unity, which would be equivalent to a point scatterer. For a distributed target scatterer the appropriate choice for $\mathbf{s}_k$ may no longer be identical to that for a point scatterer, and if the scattering vector is aspect-dependent

(which may be a good assumption for a moving target), $\mathbf{s}_k$ may change from dwell to dwell, as well. Since the complex amplitude changes induced in each multipath arrival by the scattering of the target are typically uncertain, the scattering vector, $\mathbf{s}_k$, is modeled as a zero-mean Gaussian random vector with a diagonal covariance matrix

$$\mathrm{E}[\mathbf{s}_k\mathbf{s}_q^H]=\begin{cases}\sigma_s^2\Lambda_s & \text{if } q=k,\\ \mathbf{0} & \text{if } q\neq k,\end{cases} \tag{11}$$

for all q, $k\in\{1,...,K\}$. A practical reason for this choice is that the resulting estimate should be robust to variations in the scattering coefficients. The reverberation patch scattering vector, $\mathbf{g}_{lk}$, are zero-mean Gaussian and are uncorrelated over patch index and ping index so that

$$\mathrm{E}[\mathbf{g}_{lk}\mathbf{g}_{pq}^H]=\begin{cases}\sigma_l^2\Lambda_{gl} & p=l \text{ and } q=k,\\ \mathbf{0} & p\neq l \text{ or } q\neq k,\end{cases} \tag{12}$$

for all q, $k\in\{1,...,K\}$ and p, $l\in\{1,...,L\}$ (L is the number of reverberation patches). This choice is motivated by reasons similar to those given for Eq. (11). The additive noise vector, $\mathbf{n}_k$, is also assumed to be zero-mean Gaussian with $FW\times FW$ covariance matrix

$$\mathrm{E}[\mathbf{n}_k\mathbf{n}_q^H]=\begin{cases}\sigma_n^2\mathbf{I}_n & \text{if } q=k,\\ \mathbf{0} & \text{if } q\neq k,\end{cases} \tag{13}$$

for all q, $k\in\{1,...,K\}$. It immediately follows from these assumptions that each observation, $\mathbf{x}_k$, is Gaussian and zero-mean: $\mathrm{E}[\mathbf{x}_k]=0$. Assuming that the noise and target scattering vectors are uncorrelated over the ping index, ($\mathrm{E}[\mathbf{s}_i\cdot\mathbf{n}_j^H]=\mathbf{0}$, for all i, $j\in\{1,...,K\}$), the noise and reverberation patch scattering vectors are uncorrelated over ping index ($\mathrm{E}[\mathbf{g}_{li}\cdot\mathbf{n}_j^H]=\mathbf{0}$, for all i, $j\in\{1,...,K\}$ and $l\in\{1,...L\}$), and the target scattering vectors and reverberation patch scattering vectors are uncorrelated ($\mathrm{E}[\mathbf{g}_{li}\cdot\mathbf{s}_j^H]=\mathbf{0}$, for all $i,j\in\{1,...,K\}$ and $l\in\{1,...,L\}$), then it follows that observations, $\mathbf{x}_k$, are mutually independent and that the covariance matrix for the kth observation is

$$\begin{aligned}\mathbf{R}_k(z)&=\mathrm{E}[\mathbf{x}_k\cdot\mathbf{x}_k^H|z]\\ &=\sigma_s^2\mathbf{H}_k(z)\Lambda_s\mathbf{H}_k(z)^H\\ &\quad+\sum_{i=1}^{L}\sigma_{lk}^2\mathbf{H}_{lk}(z_{lk})\Lambda_{gl}\mathbf{H}_{lk}(z_{lk})^H+\sigma_n^2\mathbf{I}.\end{aligned} \tag{14}$$

The details of how to calculate the replica matrices $\mathbf{H}_k(z)$ and $\mathbf{H}_{lk}$ are summarized by Eqs. (1) and (10). Then the pdf for the kth observation, $\mathbf{x}_k$, can be expressed as

$$p(\mathbf{x}_k|z)=\frac{1}{\pi^{FW}|\mathbf{R}_k(z)|}\exp(-\mathbf{x}_k^H\mathbf{R}_k^{-1}(z)\mathbf{x}_k). \tag{15}$$

Defining the corresponding log-likelihood as

$$L_k(z)=-\ln(\pi^{FW}|\mathbf{R}_k(z)|)-\mathbf{x}_k^H\mathbf{R}_k^{-1}(z)\mathbf{x}_k, \tag{16}$$

it follows that the ML estimate of the target depth can be written simply as

$$\hat{z}=\underset{z}{\mathrm{argmax}}\left(\sum_k L_k(z)\right). \tag{17}$$

The data input to the kth likelihood term, $L_k(z)$, in Eq. (16), is the observation, $\mathbf{x}_k$, and the modeled field is represented by $\mathbf{R}_k$. Because of the assumptions made about the natures of $\mathbf{s}_k$, $\mathbf{g}_{lk}$, and $\mathbf{n}_k$, the observations $\mathbf{x}_k$ comprise a nonstationary sequence (over ping index k) of zero-mean colored noise vectors whose covariance matrices, $\mathbf{R}_k(z)$, are given by Eq. (14). The eigenvectors and eigenvalues of $\mathbf{R}_k(z)$ are driven, in turn, by the relative delays of the multipath returns that comprise the kth ping return (as set forth in Sec. II). Thus the likelihood estimate (Eqs. (16) and (17)) uses the depth-dependent nature of the active sonar returns to localize the target in depth.

Two further details are necessary to complete the specifications of the ML estimator. The first detail is that a normalization step is performed before the above quantities are calculated: for each hypothesized depth, z, each matrix $\mathbf{H_k}(z)$ used in Eqs. (14)–(17) must be normalized by the factor $\sqrt{\mathrm{trace}(\mathbf{H_k}(z)\Lambda_s\mathbf{H_k}(z)^H)/N}$. This normalization insures that the expected total energy arriving at the receiver from the target is not a function of depth. The second detail is what estimate of the parameter σ_s^2 to use when calculating the ML estimate of the depth. Note that, although the SNR is implicitly assumed to be known, an estimate of it must be used to compute the ML estimate as set forth above. This leads to the following formula for σ_s^2, based upon expectations:

$$\sigma_s^2=\mathrm{E}[\mathbf{x}_k^H\mathbf{x}_k]/N-\sigma_n^2, \tag{18}$$

where $k\in\{1,1,...,K\}$, the depth of the target is z, and N is the length of the data vector. Unfortunately, the above expression is not directly applicable, since the expected value of the data is not available. An estimate of σ_s^2 which seems to work well in practice is

$$\widehat{\sigma_s^2}=\max_i(\mathbf{x}_i^H\mathbf{x}_i)/N-\widehat{\sigma_n^2}. \tag{19}$$

The noise level estimate, $\widehat{\sigma_n^2}$, is assumed to be obtained from the quiet periods between pings.

The maximization of the right-hand side of Eq. (17) is performed numerically. The method chosen for the simulations and real data processing (discussed in Sec. IV) is the following: choose a uniform grid of hypothesized depths that covers the water column, evaluate the log-likelihood function in Eq. (17) at each hypothesized depth on the grid, and choose the hypothesized depth at which the log-likelihood is greatest as the estimated depth of the target.

Note that, because of the assumed independence of the observations $\mathbf{x}_k$, the ML depth estimate can be efficiently updated as new data arrives at the receiver. Assume that the value of the log-likelihood at each hypothesized depth on the grid has been calculated for the available data and stored in a vector. If new data arrives later (i.e., if new pings are collected from the target), then the log-likelihood function for the new data can also be calculated for each hypothesized depth on the grid and simply added to the corresponding entry of the previously stored log-likelihood results. The updated depth estimate is the hypothesized depth at which the updated log-likelihood attains its maximum.

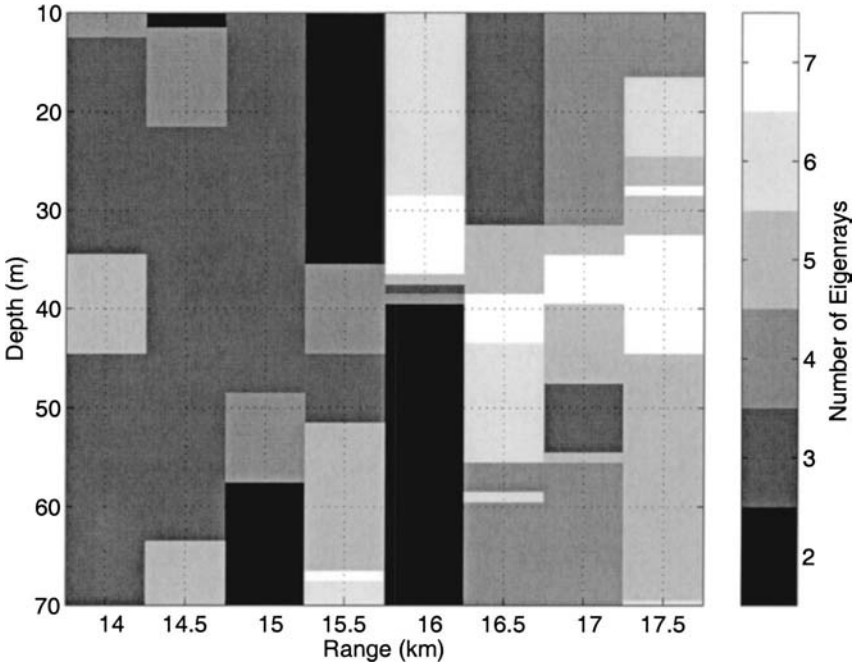


FIG. 6. This figure shows the number of rays reaching each depth at a sequence of ranges (14 to 17.5 km in 500-m increments). The number of rays is a nontrivial function of position in the range-dependent Mediterranean environment.

Finally, an important question that must be answered is how well the ML estimator can perform with this data model. To this end, the Cramér–Rao lower bound (CRLB) is of interest. The CRLB of this maximum likelihood depth estimator can be found by calculating for the Fisher information for each ping k:

$$J_k = \mathrm{Tr}\left(\mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial z} \mathbf{R}_k^{-1} \frac{\partial \mathbf{R}_k}{\partial z} \right). \tag{20}$$

This is shown, for instance, in Ref. 13. Then it follows that the CRLB for this depth estimator can be written as $(\Sigma_{k=1}^{K} J_k)^{-1}$. A comparison of this bound to the rms error of the estimates gotten from Monte Carlo trials are shown in the next section.

Note that the problem of depth estimation for active sonar is related to (but somewhat different from) target altitude estimation in the over-the-horizon (OTH) radar recently reported in Ref. 14. If the target is distant in range from the source and receiver, then the target aspect can be assumed to change slowly from one ping to the next, and so the target scattering vector, $\mathbf{s}_k$, might also change slowly over time. Under these assumptions, a Markov model for the scattering vector would be appropriate if the number of multipaths were not changing from ping to ping as the target moves. For example, this is the case for the OTH radar altitude estimation problem.[14] In the shallow-water sonar case, however, the number of multipaths frequently does change significantly with range and depth, as is illustrated in Fig. 6. If the target is moving slowly enough, the number of multipaths as a function of depth could be treated as approximately constant by ray-tracing to a baseline position which represents the average of several consecutive target positions. Under this approximation, the number of multipaths for each target position, as well as the corresponding arrival and departure angles, would remain constant as the target position changed. The multipath delays would differ from the baseline delays in accordance with the distance of the target from the baseline position and the arrival and departure angles of the various multipaths. Under these conditions, coherence of the scattering vector $\mathbf{s}_k$ over several pings might be exploited. However, Fig. 7 shows a plot of the cosine of the departure angle (at the target) versus hypothesized target depth for the various multipaths connecting the target and receiver. The environment used was that of a typical Mediterranean littoral environment. Notice that the angle cosine differences are a weak function of depth, which is unlike the OTH radar case, where these differences are increasing functions of hypothesized target altitude. As a result, the interping delay changes between multipath arrivals are not likely to yield much improvement in depth estimation performance with this environment, even when the scattering vector is perfectly corre-

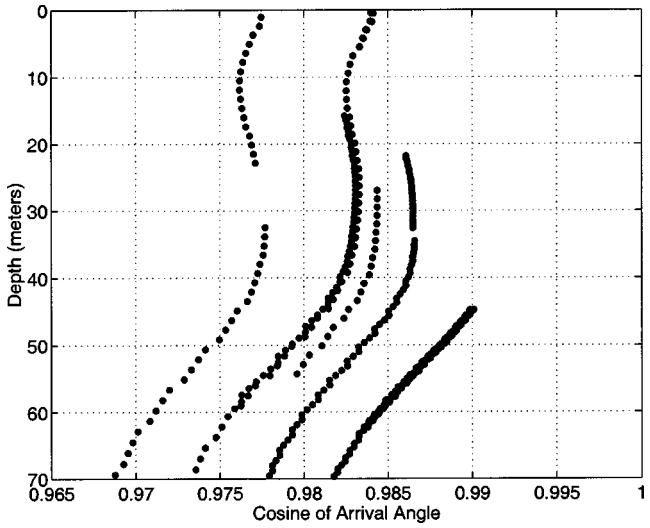


FIG. 7. A typical plot of the cosine of the departure angle (at the target) versus hypothesized target depth for the various rays connecting the target and receiver at a given ping. Note that, although the value of the angle cosine is clearly a function of depth, the depth dependence between two angle cosines at the same depth is a much weaker function of depth.

lated from ping to ping. Thus, in contrast to the matched field altitude estimation method of Ref. 14, which exploits both interdwell and intradwell modeling, in sonar only intraping modeling is used for depth estimation.

IV. SIMULATION AND REAL DATA RESULTS

In order to evaluate the performance of the ML depth estimation approach presented in this paper, Monte Carlo simulations are performed with an environmental model drawn from a Mediterranean experimental scenario. The ML depth estimation algorithm is also applied to real data collected from the same Mediterranean experiment that supplied the environmental model for the simulation performance studies. Finally, the CRLB is compared to the estimated rms errors of the Monte Carlo simulation depth estimates.

The real data were collected from an experiment performed in the littoral waters of the Mediterranean. The environment from that experiment is summarized in Fig. 2. Notice that the environment has a sloping bathymetry, and that the target is up-slope from the receiver. The data was collected by a horizontal receiver array with 144 phones (spaced 1 m apart) which was towed at a depth of 70 m. The signal waveform was from a transponder on a moving underwater platform (the target), which moved at a constant depth of 50 m. The transmitted waveform was a rectangular pulse with a duration of 0.5 s, modulated by a linear FM sweep from 395 to 400 Hz. The repetition rate of the pings was one ping every 60 s. The transponder moved monotonically from 14 to 17 km in range relative to the receiver in 15 min, and moved from about 18° to 15° (relative to the receiver array's broad side) in the same time. Although this experiment does not represent true two-way signal transmission, the depth estimation approach in this paper can still be applied by assuming that there is always exactly one source-target ray path with a transmission delay of zero seconds. Further, the transponder was effectively baffled, which ensured that its transmission pattern would correspond to an unknown scattering function (not constant over angle) as assumed for the scattering vector in this paper. As such, this data set possessed many of the pertinent features of a true two-way multipath scenario.

All phases of the implementation of the ML depth estimator were scripted in Matlab, including a ray-tracing algorithm written by the authors, which was necessary to estimate parameters such as the number of ray arrivals, arrival angles, and relative delays. The ray-tracing algorithm was based on Červený's presentation of ray theory.[15] The amount of time required to run the depth estimation code for the basic scenario of a 16-ping target track was about 2.25 min, or about 15 h for 400 consecutive runs. A two-dimensional FFT was performed to take the data from the space–time domain to the frequency-wavenumber domain. The size of the data subblock from the transformed data used for $\mathbf{x}_k$ was 5 wavenumber bins and 15 frequency bins (spaced 1.67 m^{-1} and 1.53 Hz, respectively) for a data subblock that consists of 65 frequency-wavenumber bins altogether. These parameter values are used in all of the work below, except for the study in which the bandwidth of the transmitted signal is increased: there the number of frequency bins was increased from 15 to 20 frequency bins and the number of wavenumber bins remained at 5 bins (the same frequency and wavenumber spacings were used as in the original case), which gave a total of 100 frequency-wavenumber bins in the data subblock. The number of multipaths from the target to the receiver is illustrated by Fig. 6. The number of multipaths is a nontrivial function of range and depth, and varies from two to seven paths in the environment considered here. The scattering vectors used to generate the simulated data were randomly generated with a Gaussian distribution, and were uncorrelated between pings.

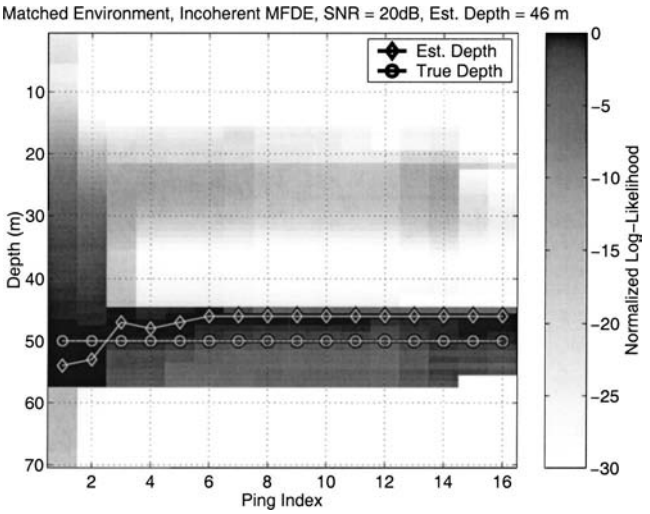


FIG. 8. Simulated log-likelihood surface evolving over 16 pings. The true depth is 50 m, and the estimated depth is 46 m after 16 pings. The SNR is 20 dB.

Figure 8 shows a simulated realization of the cumulative log-likelihood surface (or ambiguity surface), evolving over 16 pings for a SNR of 20 dB, where the SNR is defined (with reference to Eqs. (13) and (18)) as $10\log_{10}(\sigma_s^2/\sigma_n^2)$. The data was simulated using the Mediterranean environment and target-receiver tracks which were outlined earlier in this section. The ML depth estimation algorithm used the same parameters as were used for the simulated data, i.e., the as-

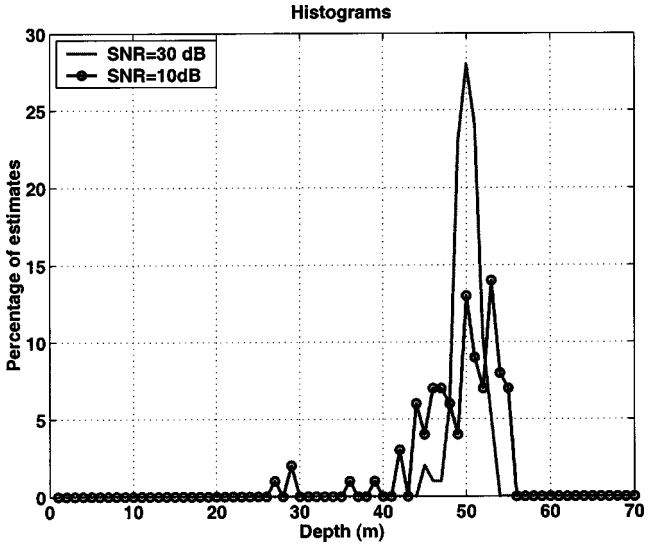


FIG. 9. Comparison of histograms of ML depth estimates at SNR values of 30 and 10 dB in the matched case.

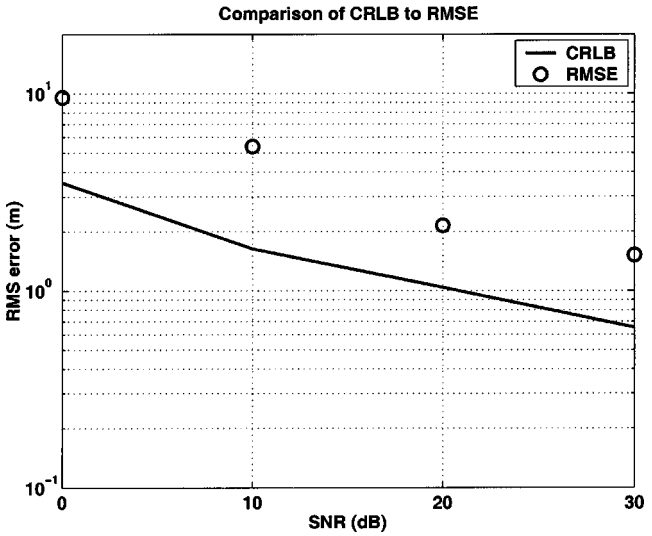


FIG. 10. Comparison of the Monte-Carlo-estimated RMSE of the depth estimate to the CRLB as a function of SNR.

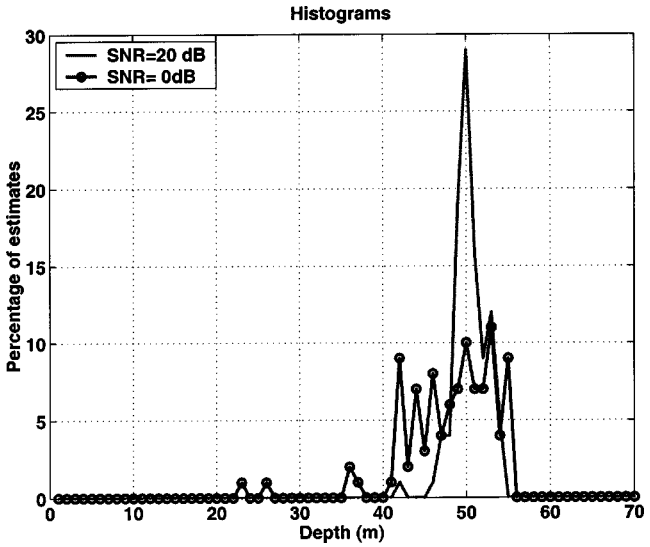


FIG. 11. Comparison of histograms of ML depth estimates at SNR values of 20 and 0 dB in the case of doubled signal bandwidth.

sumed and actual environment and tracks were matched. As each ping arrives, the new data is processed and the log-likelihood is updated in the manner prescribed by the discussion after Eq. (17). The likelihood was normalized so that the peak likelihood in each ping is zero. Notice that the width of the peak in the evolving likelihood surface rapidly decreases in width after a few pings, and that a spurious peak near the depth of 35 m diminishes also.

In order to quantify the performance of the ML estimator under known conditions, Monte Carlo simulations were performed over four different values of the SNR: 100 trials consisting of 16 pings each were simulated for SNR values 0, 10, 20, and 30 dB. For each trial at a given SNR, the changing quantities are the additive noise and the sequence of scattering vectors over the track history. The environment, target track, and receiver track used to generate the data was the same as those assumed by the depth estimation algorithms. In the sequel, this case will be referred to as the ''matched case.'' Histograms for the ML depth estimator at SNR values of 10 and 30 dB are shown in Fig. 9. As might be expected, the histogram for the 30-dB case is more sharply peaked about the target's true depth of 50 m than is the histogram for the 10-dB case. Also shown in Fig. 10, which plots the estimated root-mean-square error (RMSE) of the ML depth estimates as a function of SNR. The estimated RMSE decreases monotonically with SNR, shrinking from about 10 m at 0 dB to about 1.5 m at 30 dB.

Figure 10 shows a comparison of the CRLB to the estimated RMSE of the depth estimates from the Monte Carlo study. The bound was numerically calculated from the expression in Eq. (20), using a centered difference approximation for the depth derivatives of the covariance matrix, $\mathbf{R_k}$. For all SNR values considered here, the estimated RMSE follows the CRLB but is about 3 dB larger, which—for an SNR of 0 dB—is still only 15% of the ocean depth. This is probably because more pings need to be accumulated to insure that the ML estimator approaches the bound, as would be expected asymptotically. The ratio of the RMSE to the CRLB tended to decrease with increasing SNR.

To illustrate the effects of signal bandwidth on the performance of the ML depth estimator, Monte Carlo simulations were run with the same environment, target track, and receiver track as before, but with a transmitted signal of roughly twice the bandwidth: a rectangular pulse with a duration of 0.5 s, modulated by a linear FM sweep from 392.5 to 402.5 Hz. The number of frequency bins was increased from 15 to 20 bins, which ensures that the data subblock covers the relevant portion of the signal's bandwidth. Technically, this study also represents a matched case, but, for clarity, references to the ''matched case'' in this paper are intended to indicate the Monte Carlo study previous to this one. One hundred Monte Carlo simulations were performed over the SNR values of −10, 0, 10, and 20 dB. Notice that this sequence starts 10 dB lower than in the previous case (which ran over the values of 0, 10, 20, and 30 dB). The results are partially summarized in the comparison of histograms for the ML depth estimator at a SNR values of 20 and

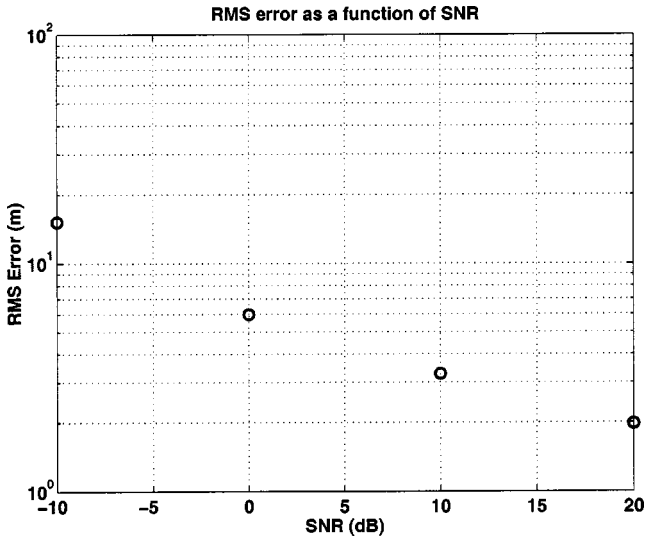


FIG. 12. Estimated RMSE of the ML depth estimates as a function of SNR for the doubled bandwidth case.

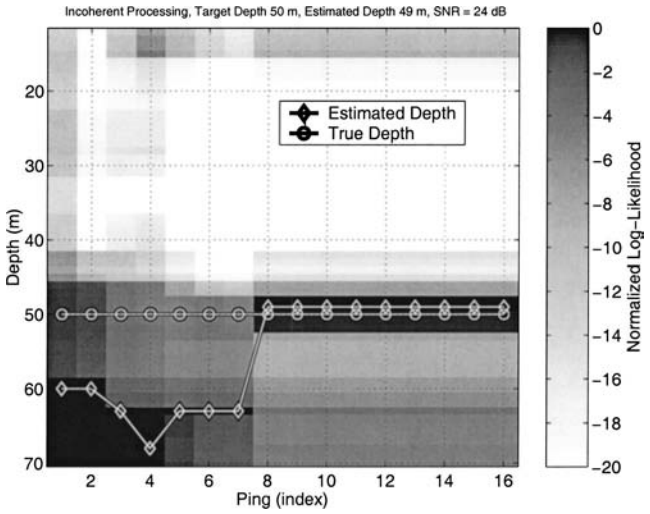


FIG. 13. Real log-likelihood surface evolving over 16 pings. The true depth of the transponder is 50 m, and the estimated depth is 49 m after 16 pings. The SNR is approximately 25 dB.

0 dB, shown in Fig. 11. Note that these histograms are similar in character to the corresponding histograms in the lower bandwidth matched case (seen in Fig. 9), but the SNR values in the present case are 10 dB lower. A further summary of the performance of the ML depth estimator is shown in Fig. 12, which plots the estimated RMSE of the ML depth estimator as a function of SNR. This curve is similar to that in Fig. 10, but with a 10-dB shift (approximately) along the SNR axis. Increasing the signal bandwidth has in this case led to an increase in the performance of the ML depth estimator.

Figure 13 shows a time-evolving cumulative log-likelihood surface for real transponder data from the Mediterranean experiment. This surface is calculated with the ML expression given in Eqs. (16) and (17). The true target depth was 50 m, and additional uncorrelated noise was added to the real data to achieve an SNR of approximately 25 dB. The ML depth estimate appears to require about eight pings to converge to within a meter of the true depth. There is a moderately strong second peak at about 60 m, and the overall level of the ambiguity surfaces is higher in the real data result than for the simulated 20 dB case shown in Fig. 8. This may be due to unresolved mismatch, but this real data result serves as a partial validation of the maximum likelihood depth estimation technique described in this paper.

V. SUMMARY AND COMMENTS

In this paper, a matched-field maximum log-likelihood estimator of target depth for active sonar was developed for multiple sonar pings. Simulations suggest that, for a Mediterranean experiment's environmental parameters and a SNR of 20 dB, a depth estimate within 10 m could be achieved after relatively few sonar pings. A Monte Carlo study was done to examine the effects of signal bandwidth on performance of these depth estimators, and it was seen that doubling the (approximately 5 Hz) bandwidth of the transmitted signal resulted in a performance increase for the ML depth estimator which was effectively equal to increasing the SNR in the original-bandwidth case by approximately 10 dB, assuming a matched environmental model. The ML depth estimator was successfully applied to real transponder data gathered in the Mediterranean experiment. The Cramér–Rao lower bound was calculated for the signal model used in this paper, and it was seen that the estimated RMSE of the ML estimator tracks the CRLB.

In conclusion, the ML depth estimator presented here shows promise whenever the relative delays between multipath arrivals are strongly depth dependent and sufficiently spread relative to the reciprocal of the signal bandwidth. Note that the ML estimator is only asymptotically optimal, and therefore it may perform badly with short data sets, particularly when the SNR is low enough for the threshold effect to dominate estimator performance. Future work will include exploration of this matched field depth estimator in a reverberation-limited scenario and studies of sensitivity to mismatch.

ACKNOWLEDGMENTS

This work was supported by ONR 32105 Contract No. N00014-99-1-0080. The authors would like to thank Jim Alsup (SPAWAR-SSC, San Diego) for his helpful discussions.

[1] H. P. Bucker, "Use of calculated sound fields and matched field detection to locate sound sources in shallow water," J. Acoust. Soc. Am. **59**, 368–373 (1976).
[2] R. G. Fizell and S. C. Wales, "Source localization in range and depth in an arctic environment," J. Acoust. Soc. Am. **78**, S57–S58 (1985).
[3] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched-field methods in ocean acoustics," IEEE J. Ocean. Eng. **18**, 401–424 (1993).
[4] J. P. Hermand and W. I. Roderick, "Acoustic model-based matched filter processing for fading time-dispersive ocean channels: Theory and experiment," IEEE J. Ocean. Eng. **18**, 447–465 (1993).
[5] H. P. Bucker, "Active matched-field tracking," J. Acoust. Soc. Am. **99**, 1783–1784 (1996).
[6] J. S. Perkins, W. A. Kuperman, L. E. Tinker, and G. T. Murphy, "Active matched-field processing," J. Acoust. Soc. Am. **91**, 2366 (1992).
[7] T. C. Yang and T. W. Yates, "Scattering from an object in a stratified medium. i. frequency dispersion and active localization," J. Acoust. Soc. Am. **96**, 1003–1019 (1994).
[8] F. Ingenito, "Scattering from an object in a stratified medium," J. Acoust. Soc. Am. **82**, 2051–2059 (1987).
[9] H. Cramér, *Mathematical Methods of Statistics* (Princeton University P., Princeton, NJ, 1963).
[10] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, Woodbury, NY, 1994).
[11] R. D. Guenther, *Modern Optics.* (Wiley, New York, 1990).
[12] J. Ward, "Space-time adaptive processing for airborne radar," Technical report, Lincoln Laboratories, M.I.T., 1994.
[13] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis* (Addison–Wesley, Reading, MA, 1995).
[14] M. Papazoglou and J. Krolik, "Matched-field estimation of aircraft altitude from multiple over-the-horizon radar revisits," IEEE Trans. Signal Process. **47**, 966–976 (1999).
[15] V. Červený, "Ray tracing algorithms in three-dimensional laterally varying layered structures," in *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*, edited by G. Nolet (Reidel, Boston, 1987), pp. 99–133.

Vibration analysis on electromagnetic-resonance-ultrasound microscopy (ERUM) for determining localized elastic constants of solids

Jiayong Tian,[a] Hirotsugu Ogi, Toyokazu Tada, and Masahiko Hirao
Graduate School of Engineering Science, Osaka University, Machikaneyama 1-3, Toyonaka, Osaka 560-8531, Japan

(Received 13 June 2003; revised 14 November 2003; accepted 24 November 2003)

In this paper we present a new acoustic-resonance microscopy, Electromagnetic-Resonance-Ultrasound Microscopy (ERUM), to measure the localized elastic stiffness of a solid material. It visualizes the resonance-frequency shift of vibrating piezoelectric crystal (langasite, $La_3Ga_5SiO_{14}$) excited by an electric field from a solenoid coil. The acoustic coupling is made only at the tip of the crystal touching the specimen surface. Being based on the calibration for the specimen's effective stiffness, the local elasticity is determined from the resonance frequencies of the crystal with the Rayleigh–Ritz method. An approximate model for the specimen's effective stiffness predicts the shift of resonance frequencies, for which the conventional Hertz-contact model is improved. As an illustrating example, the mapping of Young's modulus of a duplex stainless steel is presented, which shows good agreement with the existing study. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1642618]

PACS numbers: 43.35.Yb, 43.58.Wc [RR] Pages: 630–636

I. INTRODUCTION

Many multiphase composites are emerging and the understanding of mechanical properties of the individual phases become more important. Especially, the local elastic constants are indispensable to predict the effective elastic stiffness, the strength, and the degree of degradation of the materials.

For measuring hardness and evaluating localized elastic constants, the indentation methods are widely used.[1] Recently, ultrasonic-atomic-force microscopy (UAFM) has been developed to measure the elastic properties in micro- and nanoscale regions of materials.[2–5] It excites a flexural vibration of a microcantilever by a piezoelectric transducer. The free end of the cantilever touches a specimen surface with an applied force. The technique is capable of mapping the elasticity difference within a spatial resolution of less than 100 nm. However, much research neglected considering influence of the mounted piezoelectric transducer on the vibration of the cantilever and that of the clamping condition at the fixed end. Particularly, the clamped end is not ideally the rigid end and this uncertain boundary condition prevents one from evaluating of the elastic stiffness.

Here, we present an alternative acoustic microscopy, *Electromagnetic-Resonance-Ultrasound Microscopy* (ERUM), to quantitatively evaluate the local stiffness of a material. It uses the resonance-frequency shift of a rectangular-parallelepiped piezoelectric probe, made of langasite ($La_3Ga_5SiO_{14}$), touching the specimen only through a tip. Langasite is a piezoelectric crystal and the vibrations are excited by applying dynamic electric fields using a surrounding solenoid coil. Thus, neither an electrode nor mechanical contact is required for the acoustic coupling. Such noncontact excitation and detection of ultrasonic vibrations eliminate the measurement errors associated with the contact coupling and the ambiguous boundary condition at the supports. Such a noncontacting excitation and the detection of vibration in a piezoelectric material was first reported by Choi and Yu,[6] then by Johnson *et al.*[7] Use of the noncontacting method for scanning the object surface to measure the elastic stiffness has not been reported. In this study, we measure the resonance-frequency shift of the probe and provide an elastic-constant distribution in a localized area.

Concerning the vibration analysis, we use Lagrangian-minimization approach with the Rayleigh–Ritz approximation[8–11] to calculate the resonance-frequency shift caused by a contact with the object solid. Because the conventional Hertz model does not apply to a dynamic contact,[12] we use an approximate approach for the dynamic-contact stiffening. Lastly, we apply this new acoustic microscopy to a duplex stainless steel consisting ferritic and austenitic phases. Determined Young's moduli of the two phases were consistent with the prediction previously reported.

II. ELECTROMAGNETIC-RESONANCE-ULTRASOUND MICROSCOPY

The measurement setup of ERUM is shown in Fig. 1. An oriented rectangular-parallelepiped langasite ($La_3Ga_5SiO_{14}$) crystal stands in a solenoid coil. It measures 10.012 mm ($=L_1$) by 10.043 mm ($=L_2$) by 14.405 mm ($=L_3$), respectively. Three principal crystallographic axes x_1, x_2, and x_3 are along the three sides L_1, L_2, and L_3, respectively. The mass density is 5731 kg/m^3. A spherical bearing of tungsten

[a]Electronic mail: J_tian@me.es.osaka-u.ac.jp

 0001-4966/2004/115(2)/630/7/$20.00

carbide is bonded at the center of the bottom surface (x_3 surface), through which the langasite crystal contacts the specimen fixed on an $X-Y$ stage. A weight provides a biasing force F_0 to the nodal points on the upper surface of the probe through three needles in order to minimize the mechanical contacts with the surroundings and then enhance the sensitivity of resonance-frequency shift.

Langasite is a trigonal-symmetry crystal, whose elastic constants C_{ij}, piezoelectric coefficients e_{ij}, and dielectric coefficients ϵ_{ij} are given as[11]

$$[C_{ij}]=\begin{bmatrix} 188.5 & 104.7 & 96.87 & 14.11 & 0 & 0 \\ & 188.5 & 96.87 & -14.11 & 0 & 0 \\ & & 263.11 & 0 & 0 & 0 \\ & & & 53.35 & 0 & 0 \\ & \text{sym.} & & & 53.35 & 14.11 \\ & & & & & 41.9 \end{bmatrix} \text{(GPa)}, \tag{1}$$

$$[e_{ij}]=\begin{bmatrix} -0.429 & 0.429 & 0 & 0.193 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.193 & 0.429 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{(C/m}^2\text{)}, \tag{2}$$

$$[\epsilon_{ij}]=\begin{bmatrix} 19.05 & 0 & 0 \\ & 19.05 & 0 \\ \text{sym.} & & 50.50 \end{bmatrix} \epsilon_0, \tag{3}$$

where ϵ_0 denotes the dielectric constant in vacuum. Thus, langasite's e_{11} and e_{14} are larger in magnitude than quartz's e_{11} and e_{14} by factors 2.5 and 4.4, respectively. Such favorable piezoelectricity allows us to excite the vibration effectively without any contacts with a dynamic electric field. Furthermore, the elastic constants of langasite show a weak dependence on temperature, of the order of $10^{-5}\,\mathrm{K}^{-1}$,[13] which assures stable resonance frequencies.

We apply high-power rf bursts to the solenoid coil to cause the vibration of the probe by the converse piezoelectric effect. Then, the vibration of the probe is received by the same coil with the piezoelectric effect after the excitation.[11] The received signal is fed to a superheterodyne spectrometer to extract the signal amplitude of the same frequency component of the driving rf bursts.[14] A frequency scan provides a resonance spectrum as the one shown in Fig. 2, comprising a number of resonance peaks. The Lorentzian-fitting procedure to a peak yields the resonance frequency. We measure continuously the resonance frequency during moving the $X-Y$ stage to map the resonance-frequency shift.

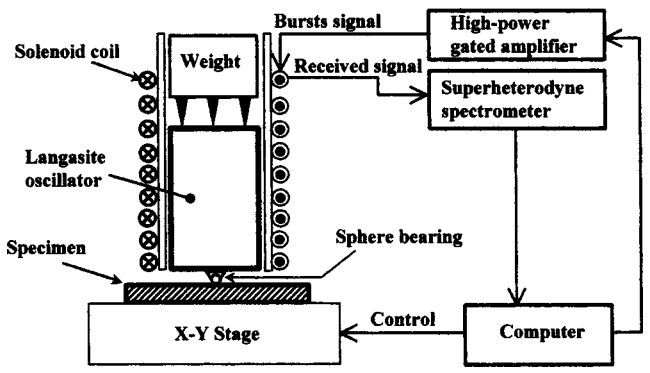


FIG. 1. Setup for ERUM.

III. VIBRATION ANALYSIS

We develop an approximated model to determine the localized Young's modulus of the specimen from the resonance-frequency shift of the probe. A simplified geometry of ERUM is shown in Fig. 3, where the tip–sample contact is equivalent to a support with springs having nonlinear spring constants k_{ij}. Here, we neglect the effect of contacts between the three needles and the crystal's upper surface because of three reasons: First, as shown later, the needles contact the nodal points of the out-of-plane displacement and affect little the out-of-plane vibrations. Second, the langasite crystal has mirror-finished surfaces, which allows almost frictionless contacts and then negligible influence on the in-plane displacement. Third, even if frictional and antinode contacts are considered, the contact areas at the needle tip are much smaller than that at the bearing–specimen contact and their influence again can be neglected, because the contact influence is enhanced with increasing the contact area, as seen in the Appendix.

The analysis contains two steps. First, we study the effect of the spring constants on the shift of the probe's resonance frequency. Second, we derive the relation between the localized Young's modulus of the specimen and the spring constants.

We adopt the Rayleigh–Ritz method for the first step. For a vibrating piezoelectric crystal with spring supports at the surface, Lagrangian Π can be expressed as[8–11]

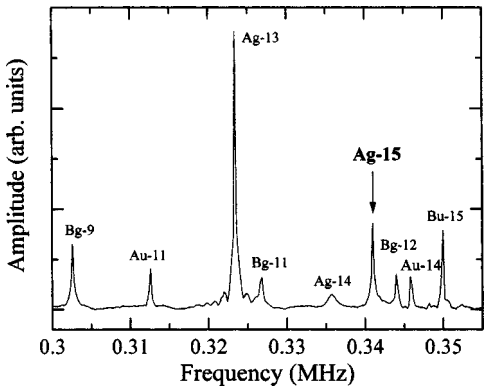


FIG. 2. Free-vibration resonance spectrum of a langasite probe measured by exiting the solenoid coil.

$$\prod = \frac{1}{2}\int_{\Omega}\int\int (S_{ij}C_{ijkl}S_{kl} + 2S_{ij}e_{ijk}\phi_{,k} - \phi_{,i}\epsilon_{ij}\phi_{,j} - \rho\omega^2 u_i u_i)dV + \frac{1}{2}\int_{\Gamma}\int \tilde{k}_{ij}u_j u_i\, dS, \tag{4}$$

where S_{ij}, ρ, u_i, and ϕ_i are the strain tensor, mass density, displacement, and electric potential, respectively. $\tilde{k}_{ij}$ are the surface distributed spring constants with dimension N/m^3. Ω and Γ denote the volume and surface of the piezoelectric crystal, respectively. We assume that $\tilde{k}_{ij}=0$ for $i\neq j$. Because no analytical solutions for u_i and ϕ exist, we approximately express the displacement vector $\mathbf{U}=[u_1,u_2,u_3]^T$ and the electric potential ϕ in the linear combination of orthogonal basis functions $\mathbf{Y}$,

$$\mathbf{U}=\mathbf{Ya}, \tag{5}$$

$$\phi=\mathbf{Yb}, \tag{6}$$

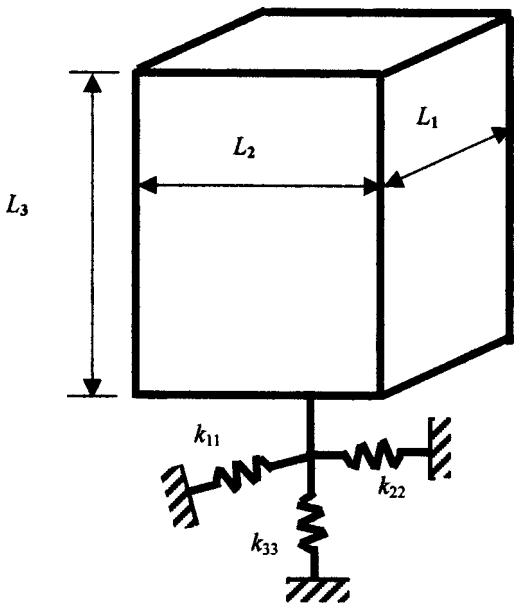


FIG. 3. Contact model of the tip.

where **a** and **b** are unknown coefficient vectors. We select Legendre polynomials P_k as the orthogonal base, which is given as

$$Y^{(p)}(x_1,x_2,x_3) = \frac{1}{\sqrt{L_1L_2L_3}}\sqrt{\frac{2k+1}{2}}\sqrt{\frac{2m+1}{2}} \times \sqrt{\frac{2n+1}{2}}P_k\left(\frac{x_1}{L_1}\right)P_m\left(\frac{x_2}{L_2}\right)P_n\left(\frac{x_3}{L_3}\right). \tag{7}$$

Note that the base functions **Y** show the orthogonality relationship:

$$\int_{\Omega}\int\int Y^{(p)}Y^{(p')}d\Omega = \delta_{kk'}\delta_{mm'}\delta_{nn'}, \tag{8}$$

where δ_{ij} denotes Kronecker's delta function. The substitution of Eqs. (5) and (6) into Eq. (4) yields

$$\prod = \frac{1}{2}\mathbf{a}^T\mathbf{K}\mathbf{a} + \mathbf{b}^T\mathbf{K}_1\mathbf{a} - \frac{1}{2}\mathbf{b}^T\mathbf{K}_2\mathbf{b} - \frac{1}{2}\omega^2\mathbf{a}^T\mathbf{M}\mathbf{a} + \frac{1}{2}\mathbf{a}^T\mathbf{K}_3\mathbf{a}, \tag{9}$$

where $K_{pp'}=\int\int\int_{\Omega}S_{ij}(Y^{(p)})C_{ijkl}S_{kl}(Y^{(p')})d\Omega$, $K_{1pp'}=\int\int\int_{\Omega}Y^{(p)}_{,i}e_{ikl}S_{kl}(Y^{(p')})d\Omega$,

$$K_{2pp'}=\int_{\Omega}\int\int Y^{(p)}_{,i}\epsilon_{ij}Y^{(p')}_{,j}\,d\Omega,$$

$$K_{3pp'}=\int_{\Gamma}\int\int Y^{(p)}_{i}\tilde{k}_{ij}Y^{(p')}_{j}dS, \quad \text{and } M_{ij}=\rho\delta_{ij}.$$

Applying the variational principle that implies the minimization of Eq. (9) with respect to **a** and **b**, we obtain

$$\mathbf{K}\mathbf{a}+\mathbf{K}_1^T\mathbf{b}+\mathbf{K}_3\mathbf{a}-\omega^2\mathbf{M}\mathbf{a}=0, \tag{10}$$

$$\mathbf{K}_2\mathbf{b}-\mathbf{K}_1\mathbf{a}=0. \tag{11}$$

Substituting Eq. (11) into Eq. (10), we obtain the characteristic equation for resonance frequencies,

$$|\mathbf{K}+\mathbf{K}_1^T\mathbf{K}_2^{-1}\mathbf{K}_1+\mathbf{K}_3-\omega^2\mathbf{M}|=0, \tag{12}$$

as well as the corresponding eigenvectors.

Langasite belongs to the trigonal system with point group 32. Free vibrations of an oriented rectangular parallelepiped of such a material fall into four groups labeled Ag, Au, Bg, and Bu,[11] whose displacement symmetry is given in Table I.

For the second step, we consider a simplification for the tip–sample contact. Figure 4 shows the detailed geometry of the tip. The tip consists of a steel base and the spherical tungsten–carbide bearing with 0.35 mm radius. The bearing can rotate, contacting the specimen. The steel base is bonded to the center of the crystal's bottom surface with the contact area of 0.6 mm radius.

The previous UFAM studies used the simple Hertz-contact model as summarized in the Appendix to calculate

TABLE I. Vibration modes of piezoelectric crystal with trigonal symmetry of elastic constants. O and E represent odd and even functions, respectively.

	k	$m+n$		k	$m+n$		k	$m+n$		k	$m+n$
u_1	O	E	u_1	E	E	u_1	E	O	u_1	O	O
u_2	E	O	u_2	O	O	u_2	O	E	u_2	E	E
u_3	E	O	u_3	O	O	u_3	O	E	u_3	E	E
ϕ	O	E	ϕ	E	E	ϕ	E	O	ϕ	O	O
	Ag			Au			Bg			Bu	

the effective contact stiffness. However, this model fails to explain the resonance-frequency shift as shown later in Fig. 7 and as discussed elsewhere.[12] We then seek a relationship between the specimen Young's modulus and resonance-frequency shift with an empirical rule as follows.

According to the Hertzian-contact theory,[15] the effective spring stiffnesses are expressed by a power law of the equivalent Young's modulus E^* or shear modulus G^*,

$$k_{33}=AE^{*\delta_1}, \tag{13}$$

$$k_{11}=k_{22}=BE^{*\delta_2}G^{*\delta_3}, \tag{14}$$

where $1/E^*=(1-\nu_1^2)/E_1+(1-\nu_2^2)/E_2$ and $1/G^*=(2-\nu_1)/G_1+(2-\nu_2)/G_2$; E, G, and ν are Young's modulus, shear modulus, and Poisson's ratio. The subscripts 1 and 2 indicate the bearing and the specimen, respectively. δ_1, δ_2, and δ_3 denote power factors. Coefficients A and B depend on the applied biasing force and shape of contacting elements. The analytical solution for contact problem is limited to a few simple cases: For example, $\delta_1=2/3$, $A=\sqrt[3]{6RF_0}$, and $\delta_2=-1/3$, $\delta_3=1$, $B=\sqrt[3]{128RF_0}$ for a normal contact of two elastic spheres without slip (a frictional contact). (When the specimen surface is well flat (mirror finish), the tangential spring constants can be negligible.) Thus, assuming that the normal spring stiffness and tangential spring stiffness show similar exponential dependences on E^* and G^* as in Eqs. (13) and (14), we determine the unknown coefficients and power factors through calibration measurements.

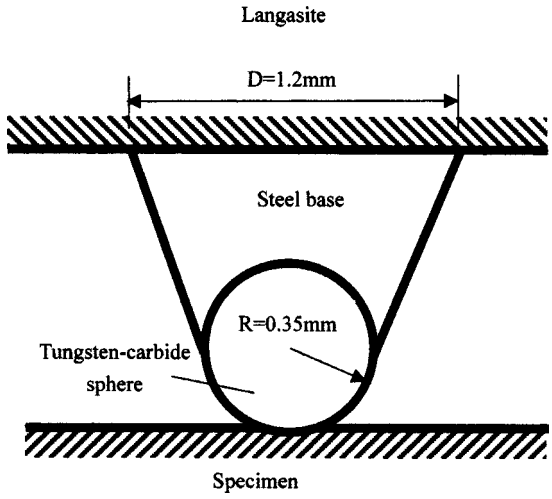


FIG. 4. Close-up of the tip for ERUM.

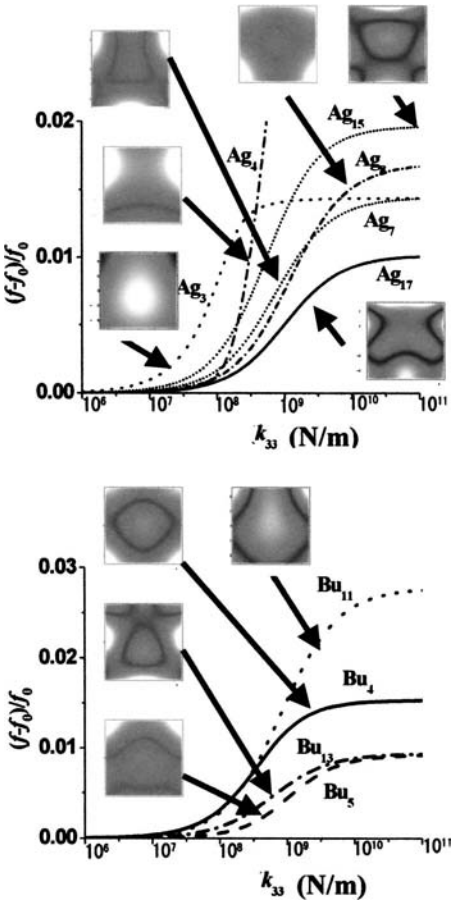


FIG. 5. Influence of the normal stiffness k_{33} on resonance frequencies for Ag and Bu groups.

IV. RESULT AND DISCUSSION

To study the effect of the normal spring constant k_{33} independently, we calculated the dependence of resonance frequencies of Ag and Bu modes on k_{33} using Eq. (12) with $k_{11}=k_{22}=0$. The result is shown in Fig. 5. Figure 5 also shows corresponding distributions of normal displacement u_3 at the probe's upper surface, which are calculated from eigenvectors associated with Eq. (12) for free vibrations. Because Au and Bg modes show a nodal line for u_3 passing the center of the bottom surface, where the tip is attached, the normal spring constant k_{33} has no influence on these modes. Shifts of resonance frequencies of Ag and Bu modes increase with increasing k_{33}. They increase with high sensitivity for $k_{33}=10^7-10^9$ N/m and show saturation with larger k_{33} values. When k_{33} is smaller than 10^8 N/m, Ag-3, Ag-15, Bu-4, and Bu-11 modes show larger sensitivity to k_{33} than others. Figure 5 predicts the largest sensitivity of Ag-3 mode to the specimen modulus (or k_{33}), but we failed to observe this mode because of the limited frequency range of the measuring instrument we used. Also, we failed to detect Bu modes with the solenoid coil we used. Hence, we use the Ag-15 mode as a most suitable mode for the present ERUM measurements because (i) this mode has relatively high sensitivity to the material stiffness (Fig. 5); (ii) it demonstrates a good spectral lineshape (see Fig. 2) and no mode overlapping occurs at its frequency; and (iii) the resonance frequency shows a small normalized temperature derivative of

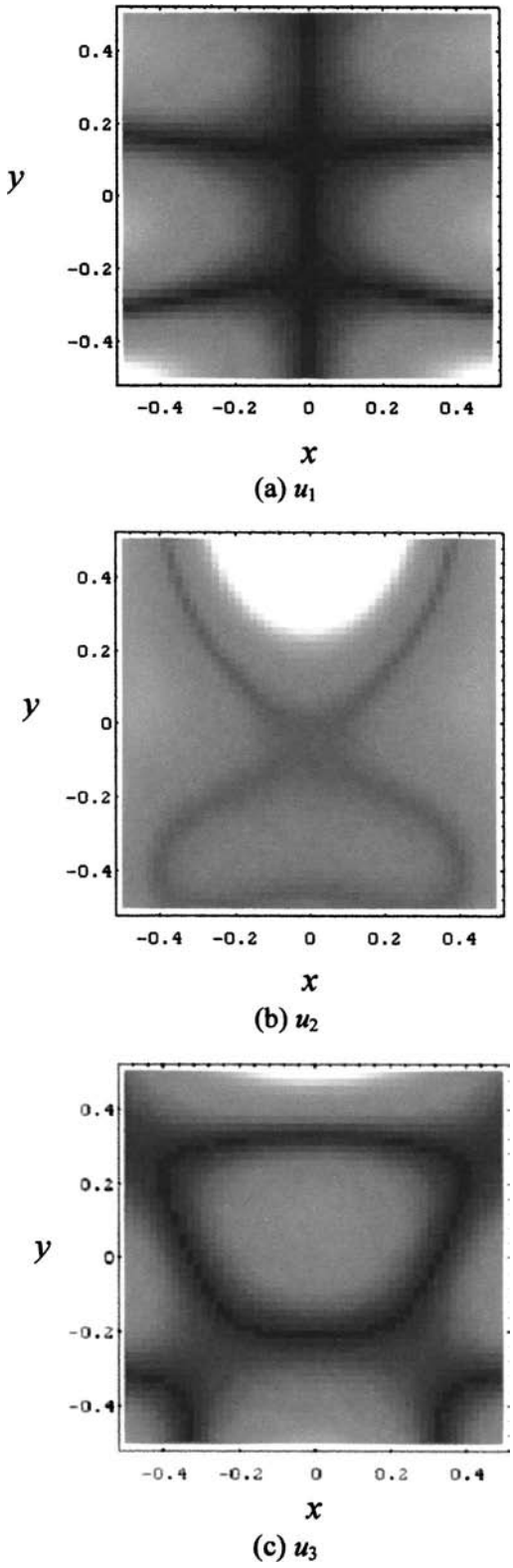


FIG. 6. The modal profile of Ag-15 resonance at the upper surface of a langasite probe.

$2.3\times10^{-5}\,\mathrm{K}^{-1}$ for the 25 °C–35 °C range, which is much smaller than that of common metals; and (iv) this mode shows nodal points for the in-plane displacements u_1 and u_2 at the center of the bottom surface, as shown in Fig. 6, which allows us to neglect the tangential spring constants k_{11} and k_{22} to make the problem much simpler. [Indeed, we calculated the effect $k_{11}(=k_{22})$ on the Ag-15 mode and verified its negligible contribution.] Thus, only the normal spring constant k_{33} affects the Ag-15 resonance frequency with unknown coefficients δ_1 and A.

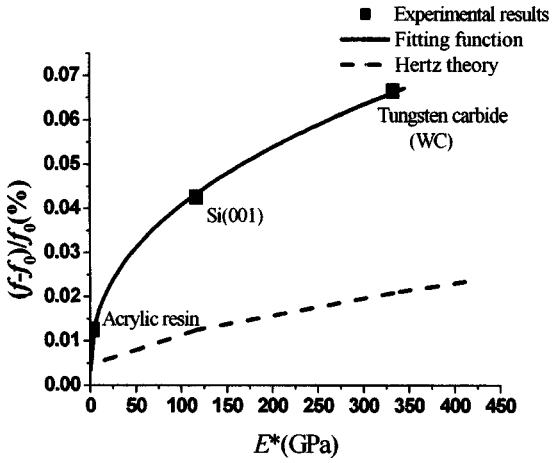


FIG. 7. A comparison between the fitting curve and Hertz theory (f_0: free vibration frequency of Ag-15).

We measured the resonance frequency of the Ag-15 mode by making the probe contact acrylic resin, (001) surface of monocrystal silicon, and polycrystalline tungsten carbide with a static force of $F_0=0.4151\,\mathrm{N}$. We then inversely determined the unknown coefficients as $A=168.12\,\mathrm{N}^{0.66}\,\mathrm{m}^{0.22}$ and $\delta_1=0.44$ by a least-square fitting with Eq. (13). Figure 7 shows the comparison between the fitting function and the measurements along with a prediction by the classical Hertz theory. The Hertz model gives a similar dependence of the frequency shift on E^*, but the magnitude is only one-fourth of that of observations, indicating that the model is inapplicable to a quantitative evaluation of specimen's modulus.

To demonstrate the applicability of our method, we obtained the elastic-constant image of a duplex stainless steel (JIS-SCS14A).[16] The material consists of 25.8% volume fraction of α (ferrite) phase and 74.2% volume fraction γ (austenite) phase. Figure 8 shows the microstructure. The γ-phase particles are precipitated in the α-phase matrix. Figure 9(a) is the ERUM image obtained by scanning the specimen surface at every 5 μm. A line trace of resonance-frequency shift is given in Fig. 9(b). We observe the resonance-frequency shift of the order of 10^{-4}. A larger shift occurs for the γ phase. Using a plausible value of 0.25 for

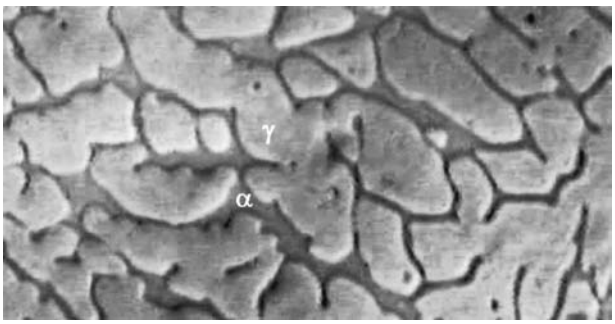


FIG. 8. Optical microstructure of a duplex stainless steel (JIS-SCS14A).

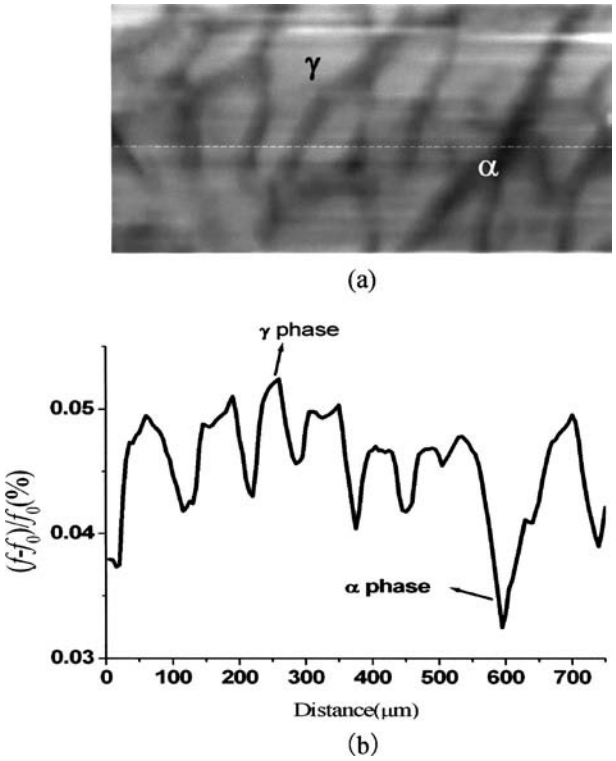


FIG. 9. (a) ERUM image of the surface of duplex stainless steel and (b) the line trace of the resonance-frequency shift, as indicated in (a).

Poisson's ratio, the fitting curve in Fig. 7 gives the grain-average Young's moduli of two phases as $E_\gamma = 219$ GPa and $E_\alpha = 128$ GPa. [Note that the specimen's Poisson ratio has an insignificant influence on the effective Young's modulus E^* because $(1-\nu^2)$ affects the modulus.] Young's modulus of γ phase agrees with that of austenitic stainless steels ($E_\gamma = 215$ GPa in Ref. 15). Young's modulus E_α is, however, considerably small compared with that of ferritic steels ($E_\alpha = 212$ GPa in Ref. 15). This is possible, considering the fact that large concentration of Cr in the α phase elongates the atomic distance. The lattice parameter of the α phase measured for the same material is 2.93 Å,[16] which is fairly larger than that of standard ferritic steels (2.87 Å). Thus, the elastic constants can reduce because of the lattice's anharmonic effect. Tane *et al.*[16] estimated the α-phase Young's modulus of the same material using a micromechanics calculation, and they gave 158 GPa, which is actually much smaller than that of standard ferritic steel by 34% and closer to our value.

V. CONCLUSION

Being based on the vibration of langasite crystal measured by the noncontacting electromagnetic-acoustic-resonance technique, ERUM shows high potential of measuring the local elastic constant of solids. The Ag-15-mode is selected for the operating mode of ERUM because of the high sensitivity to specimen modulus and the stability to temperature change. The approximate contact model is proposed by improving the basis of the conventional Hertz model. We find that the proposed approach is suitable for a quantitative evaluation of the localized modulus. The technique was applied to measure the Young-modulus distribution on a duplex stainless steel. The results show favorable agreement with the previous reports.

ACKNOWLEDGMENT

The first author (J. T.) acknowledges the support provided by Japan Society for the Promotion of Science (JSPS).

APPENDIX: HERTZ MODEL OF TIP–SAMPLE CONTACT

According to Hertz theory,[15] the contact radius a and indentation depth d of the sphere under a normal force F can be written as

$$a = \sqrt[3]{\frac{3FR}{4E^*}} \quad \text{and} \quad d = \sqrt[3]{\frac{9F^2}{16E^{*2}R}},$$

where R is the radius of the tip. The initial indentation depth d_0 can be expressed as

$$d_0 = \sqrt[3]{\frac{9F_0^2}{16E^{*2}R}}.$$

The vertical displacement at a contact point for a vibrating probe is then written as

$$u_z|_B = d - d_0 = \sqrt[3]{\frac{3F_0^2}{16E^*R}} - \sqrt[3]{\frac{3F_0^2}{16E^*R}}.$$

Because the vibration amplitudes of ERUM are of very small magnitude and F_0 is much greater than $|F-F_0|$, the vertical displacement at a contact point can be approximated to be

$$u_z|_B \approx \frac{(F-F_0)}{k_{33}},$$

by a Taylor expansion. Here $k_{33} = \sqrt[3]{6E^{*2}RF_0}$ is the effective normal-contact spring stiffness of the tip–sample contact.

If the tangential force Q_x and Q_y act and induce the elastic deformation without slip, the relative displacement between the sample and the tip can be expressed as

$$u_x = \frac{Q_x}{k_{11}} \quad \text{and} \quad u_y = \frac{Q_y}{k_{22}},$$

where $k_{11} = k_{22} = 8aG^*$ is the effective tangential-contact spring stiffness of the tip–sample contact.

[1] E. M. Uygur, "Nondestructive dynamic testing," *Research Techniques in Nondestructive Testing*, edited by R. S. Sharpe (Academic, New York, 1980), Vol. 4, pp. 205–244.

[2] K. Yamanaka and S. Nakano, "Quantitative elasticity evaluation by contact resonance in an atomic force microscopy," Appl. Phys. A: Mater. Sci. Process. **66**, S313–S317 (1998).

[3] K. Yamanaka, A. Noguchi, T. Tsuji, T. Koike, and T. Goto, "Quantitative material characterization by ultrasonic AFM," Surf. Interface Anal. **27**, 600–606 (1999).

[4] U. Rabe, J. Turner, and W. Arnold, "Analysis of the high-frequency response of atomic force microscopy cantilevers," Appl. Phys. A: Mater. Sci. Process. **66**, S277–S282 (1998).

[5] E. Kester, U. Rabe, L. Presmanes, Ph. Taihades, and W. Arnold, "Measurement of mechanical properties of nanoscaled ferrites using atomic force microscopy at ultrasonic frequencies," Nanostruct. Mater. **12**, 779–782 (1999).

[6] K. Choi and I. Yu, "Inductive detection of piezoelectric resonance by

using a pulse NMR/NQR spectrometer,'' Rev. Sci. Instrum. **60**, 3249–3252 (1989).
[7] W. Johnson, S. Kim, and D. Lauria, ''Anelastic loss in langatate,'' in *Proceedings of the IEEE/EIA International Frequency Control Symposium and Exhibition*, 2000, pp. 186–190.
[8] H. Ogi, P. Heyliger, H. Ledbetter, and S. Kim, ''Mode-selective resonance ultrasound spectroscopy of a layered parallelepiped,'' J. Acoust. Soc. Am. **108**, 2829–2834 (2000).
[9] I. Ohno, ''Rectangular parallelepiped resonance method for piezoelectric crystals and elastic constants of alpha-quartz,'' Phys. Chem. Miner. **17**, 371–378 (1990).
[10] A. Migliori, J. Sarrao, W. Visscher, T. Bell, M. Lei, Z. Fisk, and R. Leisure, ''Resonant ultrasonic spectroscopy techniques for measurement of the elastic moduli of solids,'' Physica B **183**, 1–24 (1993).
[11] H. Ogi, N. Nakamura, K. Sato, M. Hirao, and S. Uda, ''Elastic, anelastic, and piezoelectric coefficient of langasite: resonance ultrasound spectroscopy with laser-Doppler interferometry,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **50**, 553–560 (2003).
[12] J. Y. Tian, H. Ogi, and M. Hirao, ''Vibration analysis for an elastic-sphere oscillator contacting semi-infinite viscoelastic solids in resonance ultrasound microscopy,'' J. Appl. Phys., to be published.
[13] A. Bungo, C. Jian, K. Yamaguchi, Y. Sawada, S. Uda, and Y. Pisarevsky, ''Analysis of surface acoustic wave properties of the rotated Y-cut langasite substrate,'' Jpn. J. Appl. Phys. **38**, 3239–3243 (1999).
[14] M. Hirao, H. Ogi, and H. Fukuoka, ''Resonance EMAT system for acoustoelastic stress evaluation in sheet metals,'' Rev. Sci. Instrum. **64**, 3198–3205 (1993).
[15] K. L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985).
[16] M. Tane, T. Ichitsubo, H. Ogi, and M. Hirao, ''Elastic property of aged duplex stainless steel,'' Scr. Mater. **48**, 229–234 (2003).

Theoretical examination of ultrasonic pole figures via comparison with the results analyzed by finite element polycrystal model

Michiaki Kobayashi[a] and Shihua Tang
Department of Mechanical Engineering, Kitami Institute of Technology, Kitami, Hokkaido 090-8507, Japan

(Received 7 November 2003; accepted for publication 26 November 2003)

The ultrasonic wave velocities in a polycrystalline aggregate are sensitively influenced by texture development due to plastic deformation. According to Sayer's model, it is possible to construct ultrasonic pole figures via the crystallite orientation distribution function (CODF), which can be calculated by using ultrasonic wave velocities. In the previous papers, the theoretical modeling to simulate ultrasonic wave velocities propagating in solid materials under plastic deformation has been proposed by the authors and proved to be a good agreement with experimental results. Generally, wave velocities are dependent upon the propagating wave frequency; hence to evaluate texture development via ultrasonic pole figures it is necessary to examine an influence of frequency dependence on the ultrasonic wave velocities. In the present paper, the proposed theoretical modeling is applied to the texture characterization in polycrystalline aggregates of FCC metals under various plastic strain histories via ultrasonic pole figures, and also the frequency dependence is examined by using Granato–Lücke's dislocation strings model. Then the simulated ultrasonic pole figures are compared with the pole figures analyzed by the finite element polycrystal model (FEPM). The good qualitative agreement between both results suggests the sufficient accuracy of our proposed theoretical modeling.
[DOI: 10.1121/1.1642763]

PACS numbers: 43.35.Zc, 43.35.Cg, 43.25.Dc [ADP] Pages: 637–650

I. INTRODUCTION

Many alloys of common metals such as aluminum and steel are polycrystalline aggregates of single crystals having preferred orientation (texture). When a polycrystalline metal is plastically deformed, the crystallites are deformed and rotated due to the crystal slip and partially aligned to the certain crystallographic orientation called by preferred orientation or texture depending on the plastic flow geometry and magnitude. Generally, the degree and type of texture is most conveniently discussed in terms of the crystallite orientation distribution function (CODF), which gives the probability of a given crystallite in the sample having a specified orientation with respect to the sample axes. Roe[1,2] and Bunge[3] gave independent, but equivalent, methods to obtain CODF based on an expansion of CODF: $w(\xi,\psi,\phi)$ in a series of generalized spherical harmonics. The quantitative description of the ultrasonic wave velocities represented by the crystallite orientation distribution function (CODF) has been derived by Sayers[4] by adopting the Voigt's averaging method to calculate the polycrystalline elastic constants averaged over a polycrystalline aggregate of cubic crystals. More recently, a variety of such techniques has been given in several review articles and papers by Hirao *et al.*,[5] Clark *et al.*,[6] Szpunar and Hinz,[7] and Delsanto *et al.*,[8] where Hirao, Szpunar, and Clark presented nondestructive evaluation methods for predicting the anisotropy of plastic strain ratio (the Lankford value or r value), and Delsanto proposed the time-of-flight measurement techniques for the determination of both stress and texture. All of them have made a generous contribution to the research work of the nondestructive evaluation on texture with ultrasonic waves.

In the series of the authors' study, the theoretical modeling of an ultrasonic nondestructive method to evaluate plastically deformed states has been proposed and developed. Ultrasonic wave velocity changes under a uniaxial tension test and combined stress states were studied both theoretically and experimentally. The good agreement between numerical and experimental results suggested the accuracy of the proposed theoretical modeling. Successively, this method was applied to the study on the microstructural material property changes, i.e., yield surface, plastic anisotropy growth, the occurrence of the instability associated with the microslip band, and so on. The quantitative descriptions of our proposed theoretical modeling can be found in the previous papers.[9–14]

In this paper, we apply this method to characterize the texture of 6061 aluminum alloy specimens under various loading conditions of uniaxial tension, pure torsion, biaxial tension–compression, biaxial compression, and biaxial tension. The basis of this investigation lies in the relations between the orientation distribution coefficients and the elastic constants, the latter of which can be determined by ultrasonic wave velocity simulated by the proposed theoretical modeling. Then, the outlines of the theoretical procedure in this paper including the proposed theoretical modeling are briefly described in Sec. II. By the way, in general, ultrasonic wave velocity propagating in plastically deformed solids depends upon its frequency, and this probably affects ultrasonic pole figures that were determined by ultrasonic wave velocities.

[a]Electronic mail: kobayasi@mail.kitami-it.ac.jp

Therefore, we should in advance examine the frequency dependence of the acoustoplastic effect via using the Granato–Lücke model for dislocation damping[15] and discuss the effect in Sec. III. To examine the accuracy and reliability of the predicted pole figures, in Sec. IV the predicted ultrasonic results are compared with the pole figures analyzed by finite element polycrystal model (FEPM) proposed by Takahashi *et al.*[16] Finally, we proceed with the concluding summary in Sec. V.

II. THEORY

A. Outlines of proposed theoretical modeling

1. Constitutive relations

For simplicity of presentation, only rectangular Cartesian coordinates are used with the summation convention applied to repeated indices of small letter subscripts; capital letter subscripts are free indices corresponding to the small letter ones, and are not applied by the summation convention. Let a_i, X_i, and x_i denote the positions of a typical material point in the undeformed, predeformed, and current configurations, respectively. These configurations are related by the finite displacement U_i of the predeformed configuration and the additional infinite displacement u_i as follows:

$$X_i = a_i + U_i, \quad x_i = X_i + u_i. \tag{1}$$

Then, if the materials are compressible hyperelastic solids with regard to the nonlinear elastic behavior, we can define the elastic complementary energy function C as follows:

$$\begin{aligned} C &= -\frac{\lambda}{4G(3\lambda+2G)}\theta_1^2 + \frac{1}{2G}\theta_2 + \frac{1}{3}\kappa_1\theta_1^3 + \kappa_2\theta_1\theta_2 \\ &\quad + \kappa_3\theta_3, \\ \theta_1 &\equiv S_{kk}, \\ \theta_2 &\equiv \frac{1}{2}S_{ij}S_{ij}, \\ \theta_3 &\equiv \frac{1}{3}S_{ij}S_{jk}S_{ki}, \end{aligned} \tag{2}$$

in which λ and G denote Lamé constants, and S_{ij} the Piola–Kirchhoff stress tensor, respectively. The third-order elastic moduli κ_1, κ_2, and κ_3 in Eq. (2) are defined as

$$\begin{aligned} \kappa_1 &= \frac{1}{(3\lambda+2G)^3}\left\{ -l + \frac{(3\lambda+2G)^2}{4G^2}m - \frac{1}{8G^3} \right. \\ &\quad \left. \times(4G^3 + 12\lambda G^2 + 15\lambda^2 G + 6\lambda^3)n \right\}, \\ \kappa_2 &= \frac{1}{2G^3(3\lambda+2G)}\left\{ -Gm + \frac{1}{2}(\lambda+G)n \right\}, \\ \kappa_3 &= -\frac{1}{8G^3}n, \end{aligned} \tag{3}$$

where l, m, and n are the third-order Murnaghan constants.[17,18]

Based upon Eq. (2), the expressions of the Lagrangian elastic strain E^e_{ij} are obtained as follows:

$$\begin{aligned} E^e_{ij} = \frac{\partial C}{\partial S_{ij}} &= -\frac{\lambda}{2G(3\lambda+2G)}\theta_1\delta_{ij} + \frac{1}{2G}S_{ij} \\ &\quad + (\kappa_1\theta_1^2 + \kappa_2\theta_2)\delta_{ij} + \kappa_2\theta_1 S_{ij} + \kappa_3 S_{ik}S_{jk}, \end{aligned} \tag{4}$$

in which δ_{ij} is Kronecker's delta. A more detailed derivation of Eq. (3) is referred to Eq. (2)–Eq. (8) in the previous paper.[12]

If the medium is homogeneous and slightly orthotropic, we can obtain the incremental Lagrangian elastic strain $\dot{E}^e_{ij}$ as follows:

$$\begin{aligned} \dot{E}^e_{ij} &= \frac{1}{2G_{IJ}}\dot{S}_{ij} - \frac{\lambda}{2G(3\lambda+2G)}\delta_{ij}\dot{S}_{kk} + 2\delta_{ij}\kappa_1 S_{mm}\dot{S}_{kk} \\ &\quad + \kappa_2(S_{kk}\dot{S}_{ij} + S_{ij}\dot{S}_{kk} + \delta_{ij}S_{rs}\dot{S}_{rs}) \\ &\quad + \kappa_3(S_{ik}\dot{S}_{jk} + S_{jk}\dot{S}_{ik}). \end{aligned} \tag{5}$$

In Eq. (5), capital letter subscripts are free indices corresponding to the small letter ones, and are not applied by the summation convention. Here, the symmetric second-order tensor G_{IJ} is introduced to model as slightly orthotropic shear moduli due to the initial texture of the material and is expressed by the small deviation g_{ij} from isotropy as follows:

$$G_{ij} = G + g_{ij}, \quad |g_{ij}/G| \ll 1. \tag{6}$$

Let σ_{ij} denote the Cauchy stress tensor in the predeformed state; the Piola–Kirchhoff stress tensor S_{ij} is related to the Cauchy stress tensor σ_{ij} as follows:

$$S = \phi^*(\sigma), \quad S_{ij} = \frac{\rho_0}{\rho}\cdot\frac{\partial a_i}{\partial X_r}\cdot\frac{\partial a_j}{\partial X_s}\sigma_{rs}, \tag{7}$$

where ϕ^* indicates pull-back,[19] and ρ_0 and ρ are the densities of the undeformed and predeformed configurations, respectively. Here, the elastic component of the rate of deformation tensor d^e_{ij} in the predeformed configuration is related to the Lagrangian elastic strain E^e_{ij} as follows:

$$d^e = L_v\phi_*(E^e), \tag{8}$$

in which L_v is the Lie derivative and ϕ_* indicates push-forward.[19] By substituting Eqs. (5) and (7) into Eq. (8) and neglecting small terms of higher order, we can obtain the expression of d^e_{ij} as

$$d_{ij}^{e}=\left(\frac{1}{2G_{IJ}}+\frac{1}{2G}\epsilon_{kk}+\kappa_2\sigma_{kk}\right)\overset{\nabla}{\sigma}{}'_{ij}+\left(\kappa_3\sigma_{ir}-\frac{1}{G}\epsilon_{ir}\right)\overset{\nabla}{\sigma}{}'_{jr}+\left(\kappa_3\sigma_{jr}-\frac{1}{G}\epsilon_{jr}\right)\overset{\nabla}{\sigma}{}'_{ir}$$
$$+\left\{\left(\kappa_2+\frac{2}{3}\kappa_3\right)\sigma_{ij}-\frac{1}{3}\left(\frac{1}{G}+\frac{2}{G}\right)\epsilon_{ij}\right\}\overset{\nabla}{\sigma}{}_{kk}+\frac{1}{2G}\sigma_{ij}d_{kk}-\frac{1}{2G}\left(\sigma_{jr}\frac{\partial v_i}{\partial x_r}+\sigma_{ir}\frac{\partial v_j}{\partial x_r}\right)$$
$$+\frac{1}{3}\delta_{ij}\left[\left\{\frac{1}{2G_{IJ}}-\frac{3\lambda}{2G(3\lambda+2G)}+(6\kappa_1+2\kappa_2)\sigma_{mm}+\frac{1}{3}\left(\frac{1}{G}+\frac{1}{G}\right)\epsilon_{mm}\right\}\overset{\nabla}{\sigma}{}_{kk}-\left(\frac{1}{2G}-\frac{1}{G}\right)\sigma_{mm}d_{kk}\right.$$
$$\left.+\left\{3\kappa_2\sigma_{rs}+\left(\frac{1}{G}-\frac{2}{G}\right)\epsilon_{rs}\right\}\overset{\nabla}{\sigma}{}'_{rs}+\left(\frac{1}{G}-\frac{2}{G}\right)\sigma_{rs}\frac{\partial v_r}{\partial x_s}\right], \tag{9a}$$

$$\frac{1}{\overline{G}}=\frac{1}{6}\sum_{N}\frac{1}{G_{NN}}-\frac{3\lambda}{2G(3\lambda+2G)}, \tag{9b}$$

where ϵ_{ij} is the Eulerian strain-component, v_i is the velocity component of a particle, $\overset{\nabla}{\sigma}{}_{ij}$ is the Jaumann corotational stress rate of the Cauchy stress, and $\overset{\nabla}{\sigma}{}'_{ij}$ is the deviatoric component of $\overset{\nabla}{\sigma}{}_{ij}$.

If the predeformed state is the elastic or homogeneous deformation state even if including the plastic deformation, the operations of pull-back and push-forward in Eqs. (7) and (8) can be applied and we have no question in Eq. (9). However, if the predeformed state is an inhomogeneous deformation state such as the localized slip band, the operations pull-back and push-forward are no longer applicable directly. Here, considering the inhomogeneous plastic deformation due to the localized slip band to have a negligible effects on the elastic property changes of solid materials, we need to exclude the quantity of inhomogeneous plastic strain from the total strain ϵ_{ij} in Eq. (9). Therefore, considering the gradual expansion of the inhomogeneous plastic deformation area, we define $[\epsilon_{ij}^{p}]$ using the sigmoid function χ as follows:

$$\epsilon_{ij}=\epsilon_{ij}^{e}+[\epsilon_{ij}^{p}], \tag{10a}$$

$$[\epsilon_{ij}^{p}]=\int_{0}^{\epsilon_{ij}^{p}}\chi\, d\epsilon_{ij}^{p}, \tag{10b}$$

$$\chi=\frac{1}{2}\left(1-\tanh\frac{\overline{\epsilon}^{p}-\overline{\epsilon}_{th}^{p}}{\beta}\right), \tag{10c}$$

in which β is the propagating rate of the inhomogeneous plastic deformation and χ is the ratio of the component due to the homogeneous plastic deformations in the plastic strain increment, where ϵ_{ij}^{e}, ϵ_{ij}^{p} are the elastic and plastic strains, and $\overline{\epsilon}^{p}$, $\overline{\epsilon}_{th}^{p}$ are the equivalent and threshold plastic strains, respectively.

Next, we formulate the constitutive relation of the plastic response of solid materials. Considering the two types of hardening behavior, we can assume the hardening rule to be expressed as

$$\overline{\sigma}\equiv\overline{\sigma}(\overline{\epsilon}^{p},C_\nu), \tag{11a}$$

$$d\overline{\sigma}=H'\cdot d\overline{\epsilon}^{p}+K\cdot dC_\nu, \tag{11b}$$

$$H'=\left(\frac{\partial\overline{\sigma}}{\partial\overline{\epsilon}^{p}}\right)_{C_\nu=\mathrm{Const}}, \tag{11c}$$

$$K=\left(\frac{\partial\overline{\sigma}}{\partial C_\nu}\right)_{\overline{\epsilon}^{p}=\mathrm{Const}}, \tag{11d}$$

where $\overline{\sigma}$ is the equivalent stress, and C_ν is the concentration of the point defects increased by the plastic deformation. The first term on the rhs of Eq. (11b) corresponds to the conventional hardening due to line and/or plane defects such as dislocations described by the plastic potential theory in the long-range stress field, while the second term corresponds to the hardening caused by an increase of the point defects and so forth in the short-range stress field. We believe that the latter hardening is related with vertex-like yield effects induced by the texture evolution due to the growth of plastic anisotropy. Also, we assume that the increase of point defects makes a negligible contribution to plastic strains; therefore, we obtain the following restriction:

$$\left|\frac{K}{H'}\cdot\frac{dC_\nu}{d\overline{\epsilon}^{p}}\right|\ll 1. \tag{12}$$

From Eqs. (11b) and (12), the plastic deformation rate tensor d_{ij}^{p} is determined as the plastic strain rate, which is governed by the associative flow rule, and coupled strain rate d_{ij}^{c}, which is noncoaxial and is not arisen if the elastic properties are not coupled with the plastic deformation; cf. Il'lushin,[20] Dafalias,[21] and Voyiadjis:[22]

$$d_{ij}^{p}=\delta\hat{\sigma}'_{ij}d\overline{\lambda}+d_{ij}^{c}, \tag{13a}$$

$$d_{ij}^{c}=\lambda_{IJ}^{c}\overset{\nabla}{\sigma}{}'_{ij}-\frac{1}{3}\delta_{ij}\lambda_{KK}^{c}\overset{\nabla}{\sigma}{}'_{kk}, \tag{13b}$$

$$d\overline{\lambda}=\overline{\Lambda}_{rs}\overset{\nabla}{\sigma}{}'_{rs}, \tag{13c}$$

$$\overline{\Lambda}_{rs}=\frac{3\hat{\sigma}'_{rs}}{2H'\hat{\sigma}'_{kl}\hat{\sigma}'_{kl}}, \tag{13d}$$

where capital letter subscripts in Eq. (13b) are free indices corresponding to the small letter ones and λ_{IJ}^{c} is defined by

expanding coefficient K quoted in Eq. (11b) into a sixth-order symmetric tensor K_{IJklmn} to be suited to different propagating properties between transverse and longitudinal waves as follows:

$$\lambda_{IJ}^{c}=-\frac{K_{IJklmn}}{(H')^2}\cdot\frac{dC_{\nu}}{d\bar{\epsilon}^{p}}\cdot\frac{\partial\bar{\sigma}/\partial\sigma'_{kl}}{\partial\bar{\epsilon}^{p}/\partial\epsilon^{p}_{mn}}, \tag{14}$$

in which, the yield-vertex parameter $K_{IJklmn}\cdot dC_{\nu}/d\bar{\epsilon}^{p}$ is determined by experiment, and δ is a loading–unloading flag:

$$\delta=\begin{cases}0, & \text{during elastic loading or unloading,}\\ 1, & \text{during plastic loading.}\end{cases} \tag{15}$$

In Eqs. (13a)–(13d), $\overset{\nabla}{\sigma}{}'_{ij}$ is the Jaumann corotational deviatoric component of the Cauchy stress, which is introduced as the objective stress derivatives instead of the stress material derivative (see, for example, Lee[23] and Aifantis[24]), and $\hat{\sigma}'_{ij}$ is the deviatoric stress whose direction coincides with the local normal of the yield surface at the loaded point. The plastic constitutive equation described by Eq. (13) is a kind of the vertex-type or the corner theory of plasticity (see, for example, Stören and Rice,[25] Christoffersen and Hutchinson,[26] Aifantis,[24] Gotoh,[27,28] and Kuroda[29]). From Eqs. (9) and (13a), we can write the rate of deformation tensor d_{ij} as

$$d_{ij}=d^{e}_{ij}+d^{p}_{ij}. \tag{16}$$

Now we define the equivalent stress $\bar{\sigma}$ by using the non-symmetric yield function that is modified from that originally proposed by Shrivastava *et al.*[30] as follows:

$$(\tfrac{1}{3}\bar{\sigma}^2)^{3/2}=(\tfrac{1}{2}C_{ijkl}s'_{ij}s'_{kl})^{3/2}+p(\tfrac{1}{3}s'_{ij}s'_{jk}s'_{ki}), \tag{17a}$$

$$C_{ijkl}=\tfrac{1}{2}\{(\alpha_{ik}+A\cdot\epsilon^{p}_{ik})(\alpha_{jl}+A\cdot\epsilon^{p}_{jl})+(\alpha_{il}+A\cdot\epsilon^{p}_{il})(\alpha_{jk}+A\cdot\epsilon^{p}_{jk})\}, \tag{17b}$$

$$s'_{ij}=\sigma'_{ij}-m_B\epsilon^{p}_{ij}. \tag{17c}$$

Here, C_{ijkl} is plastic anisotropic modulus tensor that was originally defined by Yoshimura[31] and modified to consider the inherent anisotropy using α_{ij} instead of the Kronecker delta. A and p are anisotropic coefficients, and m_{B} Bauschinger modulus, respectively.

From Eq. (16), we can obtain the following constitutive equations:

$$\begin{aligned}\overset{\nabla}{\sigma}{}_{ij}=&2G\left(2-\frac{G}{G_{IJ}}-2G\lambda^{c}_{IJ}-\epsilon_{kk}-2G\kappa_2\sigma_{kk}\right)d'_{ij}-4G^2\delta\hat{\sigma}'_{ij}\bar{\Lambda}_{rs}d'_{rs}-4G(G\kappa_3\sigma_{ir}-\epsilon^{e}_{ir}-[\epsilon^{p}_{ir}])d'_{jr}\\&-4G(G\kappa_3\sigma_{jr}-\epsilon^{e}_{jr}-[\epsilon^{p}_{jr}])d'_{ir}-2\bar{G}\left\{G\left(\kappa_2+\frac{2}{3}\kappa_3\right)\sigma_{ij}-\frac{1}{3}\left(1+2\frac{G}{G}\right)(\epsilon^{e}_{ij}+[\epsilon^{p}_{ij}])\right\}d_{kk}-\sigma_{ij}d_{kk}+\sigma_{ir}\frac{\partial v_i}{\partial x_r}+\sigma_{ir}\frac{\partial v_j}{\partial x_r}\\&+\frac{1}{3}\delta_{ij}\bar{G}\left[\left\{1-\frac{G}{G_{IJ}}+\frac{3\lambda}{3\lambda+2G}+2\frac{G}{\bar{G}}-\frac{1}{3}\left(1+4\frac{G}{G}\right)\epsilon_{mm}-\bar{G}\left(6\kappa_1+3\kappa_2+\frac{2}{3}\kappa_3\right)\sigma_{mm}+2G\left(\kappa_2+\frac{2}{3}\kappa_3\right)\sigma_{mm}\right\}d_{kk}\right.\\&\left.+2\left\{\left(1-2\frac{G}{G}\right)(\epsilon^{e}_{rs}+[\epsilon^{p}_{rs}])-G\left(3\kappa_2+2\kappa_3-4\frac{G}{G}\kappa_3\right)\sigma_{rs}\right\}d'_{rs}+\frac{G}{G}\left(\frac{2G-\bar{G}}{G_{KK}}+4G\lambda^{c}_{KK}\right)d'_{kk}\right],\end{aligned} \tag{18}$$

where ϵ_{ij} is the Eulerian strain component, v_i the velocity component of a particle, and d'_{ij} the deviatoric component of d_{ij}, respectively. For δ, it is 1 for static plastic loading and 0 for the superposed wave motion.

2. Plane waves propagating in a principal direction

In this section, the resulting expressions for velocities of the plane waves propagating in a principal direction will be derived as an example of the generalized Christofell equation.[32]

Let the coordinate axis x_3 coincide with one of the principal directions of the stresses σ_{ij} and strains ϵ_{ij} at the pre-deformed state, and the plane wave propagate in the x_3 direction. From the generalized Christoffel equation, we obtain the following characteristic equation in terms of the matrix form:

$$\begin{vmatrix}A_{1313}-\rho V^2 & A_{1323} & 0\\ A_{2313} & A_{2323}-\rho V^2 & 0\\ 0 & 0 & A_{3333}-\rho V^2\end{vmatrix}=0, \tag{19}$$

in which the generalized Christoffel tensor A_{ijpq} is defined as

$$A_{ijpq}=2G\left(2-\frac{G}{G_{IJ}}-2G\lambda_{IJ}^{C}-\epsilon_{kk}-2G\kappa_2\sigma_{kk}\right)\left\{\frac{1}{2}(\delta_{ip}\delta_{jq}+\delta_{jp}\delta_{iq})-\frac{1}{3}\delta_{ij}\delta_{pq}\right\}-4G(G\kappa_3\sigma_{ir}-\epsilon_{ir}^{e}$$
$$-[\epsilon_{ir}^{p}])\left\{\frac{1}{2}(\delta_{jp}\delta_{rq}+\delta_{rp}\delta_{jq})-\frac{1}{3}\delta_{jr}\delta_{pq}\right\}-4G(G\kappa_3\sigma_{jr}-\epsilon_{jr}^{e}-[\epsilon_{jr}^{p}])\left\{\frac{1}{2}(\delta_{ip}\delta_{rq}+\delta_{rp}\delta_{iq})-\frac{1}{3}\delta_{ir}\delta_{pq}\right\}$$
$$-2\bar{G}\left\{G\left(\kappa_2+\frac{2}{3}\kappa_3\right)\sigma_{ij}-\frac{1}{3}\left(1+2\frac{G}{G}\right)(\epsilon_{ij}^{e}+[\epsilon_{ij}^{p}])\right\}\delta_{pq}-\sigma_{ij}\delta_{pq}+\sigma_{jr}\delta_{ip}\delta_{rq}+\sigma_{ir}\delta_{jp}\delta_{rq}$$
$$+\frac{1}{3}\delta_{ij}\bar{G}\left[\left\{1-\frac{1}{3}\left(1+4\frac{G}{G}\right)\epsilon_{kk}-\bar{G}\left(6\kappa_1+3\kappa_2+\frac{2}{3}\kappa_3\right)\sigma_{kk}+2G\left(\kappa_2+\frac{2}{3}\kappa_3\right)\sigma_{kk}\right\}\delta_{pq}\right.$$
$$+2\left\{\left(1-2\frac{G}{G}\right)(\epsilon_{rs}^{e}+[\epsilon_{rs}^{p}])-G\left(3\kappa_2+2\kappa_3-4\frac{G}{G}\kappa_3\right)\sigma_{rs}\right\}\left\{\frac{1}{2}(\delta_{rp}\delta_{sq}+\delta_{rq}\delta_{sp})-\frac{1}{3}\delta_{rs}\delta_{pq}\right\}$$
$$\left.+\frac{G}{G}\left\{\left(\frac{2G-\bar{G}}{G_{KK}}+4G\lambda_{KK}^{C}\right)\delta_{kp}\delta_{kq}-\frac{1}{3}\delta_{pq}\left(\sum_{N}\frac{2G-\bar{G}}{G_{NN}}+4G\lambda_{KK}^{C}\right)\right\}\right]$$
$$-4G^2\delta\hat{\sigma}'_{ij}\bar{\Lambda}_{rs}\left\{\frac{1}{2}(\delta_{rp}\delta_{sq}+\delta_{sp}\delta_{rq})-\frac{1}{3}\delta_{rs}\delta_{pq}\right\}. \tag{20}$$

The above expression is a three-order eigenequation related to the wave velocity V^2 propagating in the x_3 direction. By expanding this matrix equation, we get the explicit formulas for the longitudinal wave velocity V_L and the transverse wave velocities V_{T1} and V_{T2}, respectively, as follows:

$$\rho_0 V_L^2=\lambda+2G+\frac{(\lambda+2G)^2}{2G^2}(g_{11}+g_{22}+g_{33})$$
$$-2\left(1+\frac{\lambda}{G}\right)(g_{11}+g_{22})-\frac{4}{9}G^2(\lambda_{11}^{c}+\lambda_{22}^{c})$$
$$-\frac{16}{9}G^2\lambda_{33}^{c}+\left\{5+2\frac{\lambda}{G}-4G(3\lambda+2G)\left(\kappa_2+\frac{2}{3}\kappa_3\right)\right.$$
$$\left.-\frac{8}{3}G^2\kappa_3\right\}\sigma_3-\frac{1}{3}\left\{\frac{6\lambda}{3\lambda+2G}\left(\frac{\lambda}{G}+2\right)\right.$$
$$+(3\lambda+2G)^2\left(6\kappa_1+3\kappa_2+\frac{2}{3}\kappa_3\right)$$
$$\left.-4G\lambda(3\kappa_2+2\kappa_3)-\frac{8}{3}G^2\kappa_3\right\}(\sigma_1+\sigma_2+\sigma_3)$$
$$+4(\lambda+2G)[\epsilon_3^{p}], \tag{21}$$

$$\left.\begin{matrix}\rho_0 V_{T1}^2\\ \rho_0 V_{T2}^2\end{matrix}\right\}=\bar{\Gamma}+\frac{1}{2}\{g_{13}+g_{23}-2G^2(\lambda_{13}^{c}+\lambda_{23}^{c})\}$$
$$\pm\left[\Gamma^2+(\Gamma^p)^2+2\Gamma\cdot\Gamma^p\cos 2(\theta-\theta_\epsilon)+\{g_{13}-g_{23}\right.$$
$$-2G^2(\lambda_{13}^{c}-\lambda_{23}^{c})\}(\Gamma\cos 2\theta+\Gamma^p\cos 2\theta_\epsilon)$$
$$\left.+\frac{1}{4}\{g_{13}-g_{23}-2G^2(\lambda_{13}^{c}-\lambda_{23}^{c})\}^2\right]^{1/2}, \tag{22}$$

in which g_{ij} is the inherent anisotropy, ρ_0 is the mass density of the undeformed state, σ_i and ϵ_i^p are principal stress and principal plastic strain, and θ and θ_ϵ are the angles between the x_1 axis and the direction of the principal stress and the principal strain, respectively. In Eq. (22), Γ, $\bar{\Gamma}$, and Γ^p are

$$\bar{\Gamma}=G-\left\{\frac{\lambda-2G}{2(3\lambda+2G)}+G^2(2\kappa_2+\kappa_3)\right\}(\sigma_1+\sigma_2+\sigma_3)$$
$$+\left(\frac{3}{2}-G^2\kappa_3\right)\sigma_3+G[\epsilon_3^{p}],$$
$$\Gamma=\left(\frac{1}{2}-G^2\kappa_3\right)(\sigma_1-\sigma_2),\quad \Gamma^p=G([\epsilon_1^{p}]-[\epsilon_2^{p}]), \tag{23}$$

and the detailed expressions of λ_{ij}^{c} can be found in a previous paper.[33]

B. The Granato–Lücke model for dislocation damping

The strain ϵ in a solid is consisted with two kinds of strain, i.e., the elastic strain ϵ_{el} that is usually given in terms of $\epsilon_{\mathrm{el}}=\sigma/G$ by elasticity theory and, in addition, a dislocation strain ϵ_{dis} due to the motion of the dislocations under the influence of the applied stress σ, where G is the shear modulus. Consider l as the loop length of a dislocation in a cube of unit dimensions, ξ as the average displacement of length l, and Λ as the total length of movable dislocation lines, then the dislocation strain can be expressed as the function of the coordinate y on the dislocation line,

$$\epsilon_{\mathrm{dis}}=\frac{\Lambda b}{l}\int_0^l \xi(y)dy, \tag{24}$$

where b is the Burgers vector. Thus, the total strain can be clearly expressed as

$$\epsilon=\epsilon_{\mathrm{el}}+\epsilon_{\mathrm{dis}}=\frac{\sigma}{G}+\frac{\Lambda b}{l}\int_0^l \xi(y)dy. \tag{25}$$

By applying Newton's law, the equation of motion relating to the dislocation can be obtained as the form of

$$\frac{\partial^2\sigma}{\partial x^2}-\frac{\rho}{G}\frac{\partial^2\sigma}{\partial t^2}=\frac{\Lambda\rho b}{l}\frac{\partial^2}{\partial t^2}\int_0^l\xi(y)dy, \tag{26}$$

where ρ is the density of the material. On the other hand, the mathematical model for the equation of motion of a pinned down dislocation loop is given by Koehler as

$$A\frac{\partial^2\xi}{\partial t^2}+B\frac{\partial\xi}{\partial t}-C\frac{\partial^2\xi}{\partial x^2}=b\sigma, \tag{27}$$

where A is the effective mass per unit length and B is the damping force per unit length. The term C gives the force per unit length due to the effective tension in a bowed-out dislocation. A and C are defined, respectively, as

$$A=\pi\rho b^2, \quad C=\frac{2Gb^2}{\pi(1-\nu)}, \tag{28}$$

where ν is the Poisson's ratio.

Now, if we take ω as the angular frequency, v as the wave velocity, and α as the attenuation coefficient of a traveling wave in the solid, then the stress σ produced by this traveling wave can be written in the form

$$\sigma=\sigma_0\exp(-\alpha x)\left[i\omega\left(t-\frac{x}{v}\right)\right], \tag{29}$$

where σ_0 is the amplitude of stress. By taking the displacement of a dislocation as

$$\xi=\xi_0\sin\frac{\pi y}{l}\exp[i\omega t], \tag{30}$$

and substituting Eq. (29) and Eq. (30) into Eq. (26) and Eq. (27) respectively, we can obtain

$$\frac{v}{v_0}=1-\frac{4G\Lambda}{\pi^3\rho}\frac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+(\omega d)^2}, \tag{31}$$

where v_0 is the wave velocity at the original state (undeformed configuration), ω_0 is the resonant frequency, and d is the resistance coefficient, respectively. v_0, ω_0, and d are defined as

$$v_0^2=\frac{G}{\rho}, \quad \omega_0^2=\frac{2G}{\rho(1-\nu)}\cdot\frac{1}{l^2}, \quad d=\frac{B}{A}. \tag{32}$$

For the investigation of the frequency dependence of the acoustoplastic effect, it requires much information of ultrasonic wave velocities at a variety of frequencies. However, it is rather laborious and time consuming by only experimental research work. Therefore, in the present paper, we try to make an evaluation on the frequency dependence of the acoustoplastic effect by numerical simulations based upon the required minimum experimental results. If we take $(v/v_0)_{\omega_p}$ as the velocity changes at a certain given frequency ω_p, then the velocity changes $(v/v_0)_\omega$ at an arbitrary frequency ω can be derived from Eq. (31). Clearly

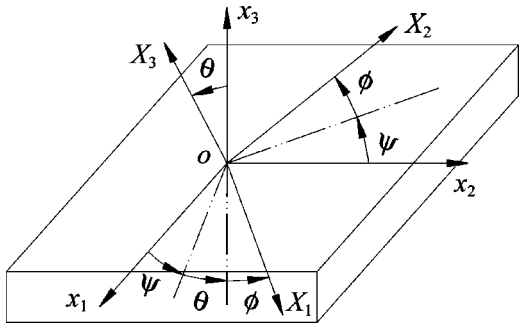


FIG. 1. Coordinate systems of sample and crystals.

$$\left(\frac{v}{v_0}\right)_\omega=1-\frac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+(\omega d)^2}\times\frac{(\omega_0^2-\omega_p^2)^2+(\omega_p d)^2}{\omega_0^2-\omega_p^2}\cdot\left\{1-\left(\frac{v}{v_0}\right)_{\omega_p}\right\}, \tag{33}$$

where ω_0 and d can be determined by ultrasonic wave velocity measurements at three different frequencies, as illustrated in Eq. (31). Here, it should be emphasized that from Eq. (31) the following notable relations can be derived: wave velocity in the adiabatic process, e.g., the impact response, is obtained by ω to be infinity:

$$\frac{v}{v_0}=\lim_{\omega\to\infty}\left\{1-\frac{4G\Lambda}{\pi^3\rho}\cdot\frac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+(\omega d)^2}\right\}=1; \tag{34a}$$

wave velocity in an isothermal process, e.g., the static response, is obtained by ω to be zero:

$$\frac{v}{v_0}=\lim_{\omega\to0}\left\{1-\frac{4G\Lambda}{\pi^3\rho}\cdot\frac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+(\omega d)^2}\right\}=1-\frac{4G\Lambda}{\pi^3\rho\omega_0^2}. \tag{34b}$$

Therefore, Eq. (34) suggests that due to the movable dislocation lines the wave velocity does not change in the adiabatic process and change in the isothermal process, respectively.

C. Ultrasonic waves and ultrasonic pole figures in a textured material

1. The crystallite orientation distribution function

The qualitative description of the ultrasonic wave velocities propagating in the sample requires knowledge of the orientation distribution of crystallites in the sample. Let $o-x_1x_2x_3$ be an orthogonal set of reference axes fixed in the sample with x_1 and x_2 on the plate surface, and x_3 be the wave propagating direction. Let $o-X_1X_2X_3$ be an orthogonal set of axes for a crystallite given by the $\langle100\rangle$, $\langle010\rangle$, and $\langle001\rangle$ crystallographic directions. The orientation of a given crystallite with respect to the sample coordinate system may then be defined uniquely by the three Euler angles θ, ψ and ϕ, as shown in Fig. 1. The crystal coordinate X_i is obtained from the sample coordinate x_i by the following: (i) a rotation of ψ about ox_3,

TABLE I. Components of transformation matrix R_{ij}.

	x_1	x_2	x_3
X_1	$\cos\theta\cos\psi\cos\phi-\sin\psi\sin\phi$	$-\cos\theta\cos\psi\sin\phi-\sin\psi\cos\phi$	$\sin\theta\cos\psi$
X_2	$\cos\theta\sin\psi\cos\phi+\cos\psi\sin\phi$	$-\cos\theta\sin\psi\sin\phi+\cos\psi\cos\phi$	$\sin\theta\sin\psi$
X_3	$-\sin\theta\cos\phi$	$\sin\theta\sin\phi$	$\cos\theta$

$$x_1'=x_1\cos\psi+x_2\sin\psi,$$

$$x_2'=-x_1\sin\psi+x_2\cos\psi,\quad x_3'=x_3; \tag{35}$$

(ii) a rotation of θ about ox_2',

$$X_1''=x_1'\cos\theta-x_3'\sin\theta,\quad X_2''=x_2',$$

$$X_3''=x_1'\sin\theta+x_3'\cos\theta; \tag{36}$$

(iii) a rotation of ϕ about oX_3,

$$X_1=X_1''\cos\phi+X_2''\sin\phi,$$

$$X_2=-X_1''\sin\phi+X_2''\cos\phi,\quad X_3=X_3''. \tag{37}$$

The combination of Eqs. (35), (36), and (37) gives

$$\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix}=\begin{bmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\cos\theta & 0 & -\sin\theta\\ 0 & 1 & 0\\ \sin\theta & 0 & \cos\theta\end{bmatrix}\times\begin{bmatrix}\cos\psi & \sin\psi & 0\\ -\sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}. \tag{38}$$

It can be rewritten as

$$X_i=R_{ij}x_j\quad\text{or}\quad x_i=R_{ji}X_j, \tag{39}$$

where R_{ij} is called the transformation matrix, and its components are listed in Table I.

The orientation distribution of crystallites can be represented by the crystallite orientation distribution function (CODF) $w(\xi,\psi,\phi)$, where $\xi=\cos\theta$. It is convenient to expand CODF as a series. Clearly,

$$w(\xi,\psi,\phi)=\sum_{l=0}^{\infty}\sum_{m=-1}^{1}\sum_{n=-1}^{1}W_{lmn}P_n^l(\xi)\times\exp(-im\psi)\exp(-in\phi), \tag{40}$$

where $P_n^l(\xi)$ are the generalized associated Legendre function, and W_{lmn} are the orientation distribution coefficients (ODCs), respectively.

2. Ultrasonic waves in a textured material

The relations between the elastic constants and ODCs based upon the Voigt averaging methods have been developed by Sayers,[4] together with the explicit expressions of ultrasonic wave velocities in anisotropic polycrystalline aggregates in terms of ODCs and are reviewed here briefly.

To simplify the problems, it is assumed that the densities of all crystallites in the sample are equal so that the calculation of the ultrasonic wave velocities involves an average over only the single crystal elastic constants. It is also assumed the polycrystalline aggregate has orthorhombic symmetry, i.e., it is assumed to possess three orthogonal mirror planes given by the planes x_1x_2, x_2x_3, and x_3x_1, for example, a rolled plate. Take a_{ij} as the direction cosines of the crystal frame with respect to the sample frame; then the elastic constants C'_{ijkl} in the sample frame can be expressed in terms of those in the crystal frame C_{ijkl} by

$$C'_{ijkl}=a_{pi}a_{qj}a_{rk}a_{sl}C_{ijkl}, \tag{41}$$

where a_{ij} take their values the same as R_{ij}, as listed in Table I.

Adopt the approach of Voigt for averaging the elastic constants and perform the integral of Eq. (41) over Euler angles to yield the clear expressions of C'_{ijkl} as

$$\begin{aligned}
\langle C'_{11}\rangle&=C_{11}-\tfrac{2}{5}C_\nu-2C_\nu\delta_1,\\
\langle C'_{22}\rangle&=C_{11}-\tfrac{2}{5}C_\nu-2C_\nu\delta_2,\\
\langle C'_{33}\rangle&=C_{11}-\tfrac{2}{5}C_\nu-2C_\nu\delta_3,\\
\langle C'_{44}\rangle&=C_{44}+\tfrac{1}{5}C_\nu+C_\nu\delta_4,\\
\langle C'_{55}\rangle&=C_{44}+\tfrac{1}{5}C_\nu+C_\nu\delta_5,\\
\langle C'_{66}\rangle&=C_{44}+\tfrac{1}{5}C_\nu+C_\nu\delta_6,\\
\langle C'_{23}\rangle&=C_{12}+\tfrac{1}{5}C_\nu+C_\nu\delta_4,\\
\langle C'_{31}\rangle&=C_{12}+\tfrac{1}{5}C_\nu+C_\nu\delta_5,\\
\langle C'_{12}\rangle&=C_{12}+\tfrac{1}{5}C_\nu+C_\nu\delta_6,\\
C_\nu&=C_{11}-C_{12}-2C_{44}.
\end{aligned} \tag{42}$$

It shall be pointed out here that it is always convenient to take advantage of the symmetries to represent the fourth-rank tensor C_{ijkl} (or C'_{ijkl}) as a symmetric 6×6 matrix. To convert C_{ijkl} (or C'_{ijkl}) into a two-dimensional matrix, we adopt the convention in Nye[34] and replace the six distinct pairs ij with a single integer as follows:

Four index tensor indices

ij 11 22 33 23,32 31,13 12,21

Two index matrix indices

m 1 2 3 4 5 6.

The angular brackets in Eq. (42) indicate an average over all the crystallites in the sample, i.e., an average over the crystalline orientation distribution function $w(\xi,\psi,\phi)$. The symbols δ_1, $\delta_2,\ldots,\delta_6$ in Eq. (42) take the following forms:

$$
\begin{aligned}
\delta_1 &= -\frac{6\sqrt{2}}{35}\pi^2\left(W_{400}-\frac{2\sqrt{10}}{3}W_{420}+\frac{\sqrt{70}}{3}W_{440}\right),\\
\delta_2 &= -\frac{6\sqrt{2}}{35}\pi^2\left(W_{400}+\frac{2\sqrt{10}}{3}W_{420}+\frac{\sqrt{70}}{3}W_{440}\right),\\
\delta_3 &= -\frac{16\sqrt{2}}{35}\pi^2 W_{400},\\
\delta_4 &= -\frac{16\sqrt{2}}{35}\pi^2\left(W_{400}+\sqrt{\frac{5}{2}}W_{420}\right),\\
\delta_5 &= -\frac{16\sqrt{2}}{35}\pi^2\left(W_{400}-\sqrt{\frac{5}{2}}W_{420}\right),\\
\delta_6 &= \frac{4\sqrt{2}}{35}\pi^2(W_{400}-\sqrt{70}W_{420}).
\end{aligned}
\tag{43}
$$

Let the coordinate axes of sample coincide with the directions of principal stresses. For a plane wave propagating in the direction of the x_3 axis at the predeformed state, the longitudinal wave velocity V_L^C and the transverse wave velocities V_{T1}^C, V_{T2}^C are

$$
V_L^C=\sqrt{\frac{\langle C'_{33}\rangle}{\rho_0}},\quad V_{T1}^C=\sqrt{\frac{\langle C'_{55}\rangle}{\rho_0}},\quad V_{T2}^C=\sqrt{\frac{\langle C'_{44}\rangle}{\rho_0}},
\tag{44}
$$

where ρ_0 is the density of the sample. If the velocities V_{L0} and V_{T0} in a natural state are expressed in terms of Lamé constants λ and G as $V_{L0}^2=(\lambda+2G)/\rho_0$ and $V_{T0}^2=G/\rho_0$, respectively, then the explicit formulas for wave velocities propagating in a textured material can be obtainable by combining Eqs. (42)–(44). Clearly

$$
\begin{aligned}
\frac{V_L^C}{V_{L0}} &= 1+\frac{16\sqrt{2}\pi^2}{35(\lambda+2G)}C_\nu W_{400},\\
\frac{V_{T1}^C}{V_{T0}} &= 1-\frac{8\sqrt{2}\pi^2}{35G}C_\nu\left(W_{400}-\sqrt{\frac{5}{2}}W_{420}\right),\\
\frac{V_{T2}^C}{V_{T0}} &= 1-\frac{8\sqrt{2}\pi^2}{35G}C_\nu\left(W_{400}+\sqrt{\frac{5}{2}}W_{420}\right).
\end{aligned}
\tag{45}
$$

As shown in Eq. (45), the substitution of the orientation distribution coefficients W_{400} and W_{420} into Eq. (45) derives longitudinal and transverse wave velocities in textured materials, that is, the ultrasonic wave velocities are dependent upon the ODCs in textured materials. Therefore, it is possible to predict the texture behaviors from ultrasonic wave velocities conversely.

3. Ultrasonic pole figures

Let t specify the normal to a given crystallographic plane, χ and η denote polar and azimuthal angles between t and axes of sample coordinate system $o-x_1x_2x_3$, as shown in Fig. 2. Roe[2] gave the relation between the ODCs and the normalized pole intensity $q(\zeta,\eta)(\zeta=\cos\chi)$, that is,

$$
q(\zeta,\eta)=\frac{1}{4\pi}+\sum_{l=4}^{\infty}\left\{S_l\sum_{m=0}^{1}P_l^m(\zeta)W_{lm0}\cos m\eta\right\},
\tag{46}
$$

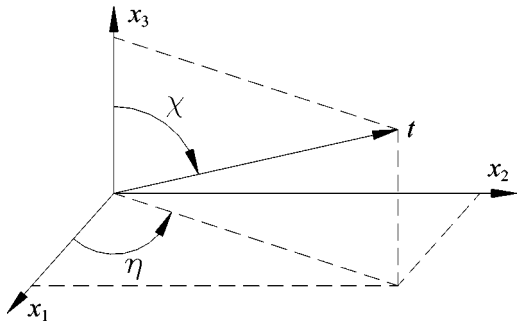


FIG. 2. Orientation of plane normal t with respect to the sample coordinate system $o-x_1x_2x_3$.

where $P_l^m(\zeta)$ is the normalized associated Legendre function. S_l is determined as follows according to different pole figures:

$$
\begin{aligned}
&\langle 100\rangle \text{ pole figure } S_4=2\pi,\quad S_6=2\pi,\ldots,\\
&\langle 110\rangle \text{ pole figure } S_4=-\pi/2,\quad S_6=-13\pi/4,\ldots,\\
&\langle 111\rangle \text{ pole figure } S_4=-4\pi/3,\quad S_6=32\pi/9,\ldots.
\end{aligned}
\tag{47}
$$

In the present work, only transverse waves are used to determine W_{400} and W_{420} for a calculation of the normalized pole density $q(\zeta,\eta)$ in the $\langle 111\rangle$ pole figure; consequently, Eq. (46) is simplified to

$$
\begin{aligned}
q(\zeta,\eta)=\frac{1}{4\pi}-\frac{4\pi}{3}\Big\{&\frac{1}{8}(35\zeta^4-30\zeta^2+3)W_{400}\\
&-7.5(7\zeta^4-8\zeta^2+1)W_{420}\cos(2\eta)\Big\}.
\end{aligned}
\tag{48}
$$

Here, the texture is usually discussed in terms of the ultrasonic pole figure, which is a stereographic projection of the normalized pole density onto a convenient plane. The stereographic projection method will be introduced in the next section.

III. EXAMINATION OF FREQENCY DEPENDENCE OF ACOUSTOPLASTIC EFFECT AND TEXTUTRE PREDICTION VIA ULTRASONIC POLE FIGURES

A. Examination of frequency dependence of acoustoplastic effect

Generally, propagating characteristics of an ultrasonic wave are not only dependent upon the microstructural properties of the plastically deformed materials, but also dependent upon the propagating wave frequencies. Therefore, to accomplish a precise evaluation of material property changes, it is necessary to examine the influence of frequency dependence on the propagating wave velocities. Here, we take aluminum 6061-T6 as the study object, and the internal state variables used for numerical simulations are those determined by ultrasonic wave velocity measurements with 5 MHz central frequency under a uniaxial tension test,[35] as listed in Table II.

As illustrated by Eq. (31), velocity measurements at three different frequencies are required to determine the total length of movable dislocation line Λ, resonant frequency ω_0, and resistance coefficient d. Therefore, the longitudinal

TABLE II. Internal state variables used in numerical simulations.

Anisotropic coefficients	$A=-3.2\overline{\nu}^{a}$
	$p=7.0\overline{\epsilon^p}\,\overline{\nu}$
	$\alpha_{ij}=\delta_{ij}{}^{b}$
Bauschinger modulus	$m_B=350(\overline{\epsilon^p})^{0.07}$ MPa
Yield-vertex parameter	$K_{ijklmn}\cdot dC_\nu/d\overline{\epsilon}$
	$=\{\delta_{ij}^{c}(28\overline{\epsilon^p}-0.09/\overline{\epsilon^p})\exp(-20\overline{\epsilon^p})+\delta_{ij}(2.5-34\overline{\epsilon^p})(\overline{\epsilon^p})^{-0.22}\}$
	$\times\delta_{ik}\delta_{km}\delta_{jl}\delta_{lm}$ MPa (i, j, k, l not summed)
Threshold plastic strain	$\overline{\epsilon_{th}^p}=2.6\%$

[a] $\overline{\nu}=(2\epsilon_3^p-\epsilon_1^p-\epsilon_2^p)/(\epsilon_1^p-\epsilon_2^p)$.
[b] δ_{ij} : Kronecker delta.
[c] $\delta_{ij}=0$, if $i=j$; 1, if $i\neq j$.

wave velocity changes under the uniaxial tension test are measured at frequencies of 5, 10, and 200 MHz, as denoted by solid circles, triangles, and rectangles in Fig. 3, respectively. In case of velocity changes at frequencies of 5 and 10 MHz, which are measured by a sing-around method. As concerned with longitudinal, or leaky creeping wave velocity changes at a frequency of 200 MHz, they are determined from $V(z)$ curve measurements by using a point-focused scanning acoustic microscope (SAM). Based upon these measured wave velocities at three different frequencies of 5, 10, and 200 MHz shown in Fig. 3, we can calculate Λ, ω_0, and d via Eq. (31). The solid lines in Fig. 3 represent the numerical simulations of longitudinal wave velocity changes at the range of frequency from 1 to 200 MHz via Eq. (33) and the above calculated values of Λ, ω_0, and d.

From Fig. 3 we can see that the discrepancy of longitudinal wave velocities between the different frequencies becomes smaller and smaller as the frequency decreases; they are nearly overlapped with each other from frequencies of 1–5 MHz, namely, at very low frequency ranges (less than 5 MHz) the ultrasonic wave velocity is independent upon its frequency. On the other hand, velocity changes are almost negligible for both experimental and numerical results at a frequency of 200 MHz; this indicates that the effect on longitudinal wave velocity changes due to dislocation damping can be ignored at very high frequency ranges. The foregoing characteristics of ultrasonic wave velocity are clearly identical to the suggestions from Eq. (34) in Sec. II B.

It seems from the preceding analysis that it is reasonable to construct ultrasonic pole figures at the frequency range of less than 5 MHz, because it shows no frequency dependence of ultrasonic wave velocities. Otherwise, the pattern of ultrasonic pole figures are probably changed from one to another

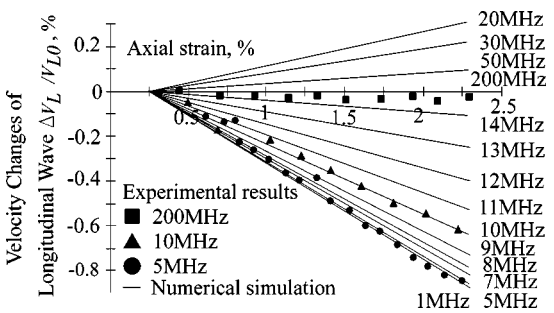


FIG. 3. Frequency dependence of longitudinal wave velocity in the plastic deformation range.

at the range of 5 to 200 MHz according to the frequency dependence of the wave velocity, as indicated by Eqs. (45) and (48). Therefore, in the next subsection ultrasonic pole figures are constructed using transverse wave velocities at a frequency of 5 MHz.

B. Ultrasonic pole figures

1. Transverse wave velocity changes

In this section, the proposed theoretical modeling of an ultrasonic nondestructive evaluation method is applied to simulate the transverse wave velocity changes under various loading conditions of uniaxial tension ($\sigma_1:\sigma_2:\sigma_3=1:0:0$), pure torsion (simple shear), biaxial tension–compression (pure shear; $\sigma_1:\sigma_2:\sigma_3=1:-1:0$), biaxial compression ($\sigma_1:\sigma_2:\sigma_3=-1:-1:0$), and biaxial tension ($\sigma_1:\sigma_2:\sigma_3=1:1:0$). The transverse wave velocities simulated by the proposed theoretical modeling are presented with solid lines in Figs. 4(a)–4(c), respectively. Here, the ultrasonic wave is supposed to propagate along the x_3 direction. Two modes of transverse waves V_{T1} and V_{T2} in Figs. 4(a)–4(c) denote the velocities polarized along the principal directions of stress, respectively, and caused only by plastic deformation without the acousoelastic effect.

It can be seen from Figs. 4(a)–4(b) that velocity changes polarized along the parallel (V_{T1}) and perpendicular (V_{T2}) to a principal direction of stress are different from each other as the development of plastic deformation under uniaxial tension, pure torsion, and biaxial tension–compression. However, velocity changes of V_{T1} and V_{T2} are overlapped with each other under biaxial compression, so for that under biaxial tension, as shown in Fig. 4(c). Transverse wave velocity changes under simple and pure shear states illustrated in Fig. 4(b) show nearly the same tendencies as the development of plastic deformation.

2. Ultrasonic pole figures

Here, in order to avoid verbosity, the uniaxial tension test will be taken as an example to explain how the ultrasonic pole figure is constructed.

First, based upon the numerical results of transverse wave velocities simulated by the proposed theoretical modeling under a uniaxial tension test shown in Fig. 4(a), the orientation distribution coefficients W_{400} and W_{420} are calculated via Eq. (45), as shown in Fig. 5. Here, the hollow and solid circles represent W_{400} and W_{420}, respectively. Next, the

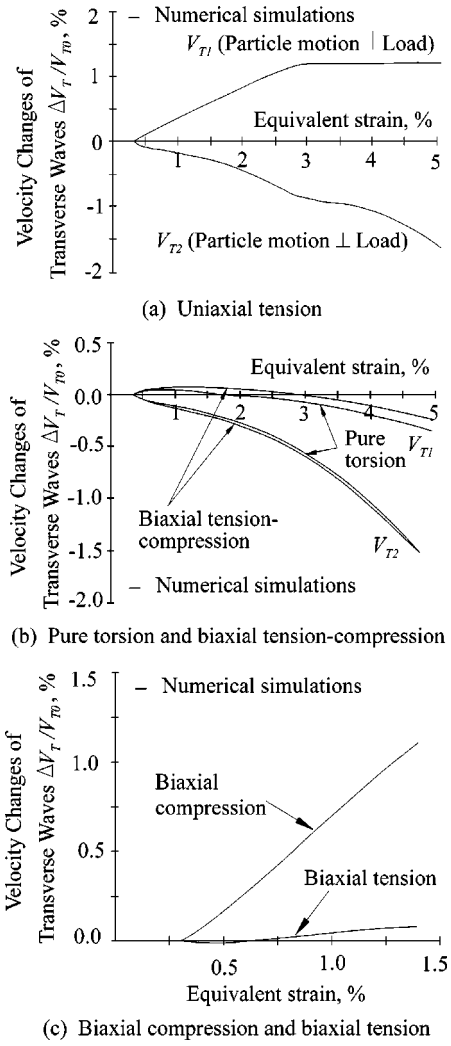


FIG. 4. Velocity changes of transverse waves under uniaxial tension, pure torsion, biaxial tension–compression, biaxial compression, and biaxial tension in the plastic deformation range.

normalized pole intensity $q(\zeta,\eta)$ is calculated from the determined ODCs W_{400} and W_{420} at one certain plastic deformation point (for example, at equivalent plastic strain $\overline{\epsilon^p}=4.98\%$). Thus, the {111} ultrasonic pole figure can be constructed by projecting the calculated normalized pole density onto the projection plane for the ⟨111⟩ crystallographic direction, as shown in Fig. 6(a).

For the same way, the {111} ultrasonic pole figures under the other loading conditions of pure torsion, biaxial tension–compression, biaxial compression, and biaxial tension can also be constructed, as shown in Figs. 6(b)–6(e), respectively. It should be pointed out here that the longitudinal wave velocity change due to the plastic deformation under pure torsion (simple shear state) is deduced to be zero via theoretical consideration.[35] So it can be drawn from Eq. (45) that the orientation distribution coefficient W_{400} is also zero. However, velocity changes of the longitudinal wave under biaxial tension–compression (pure shear state) are apparently different from zero,[35,36] so W_{400} is also different from zero according to Eq. (45). This difference of W_{400} under simple and pure shear states is responsible for the difference of the ultrasonic pole figures near the central area under both states. However, not so much of a remarkable difference of ultrasonic pole figures under simple and pure shear states is

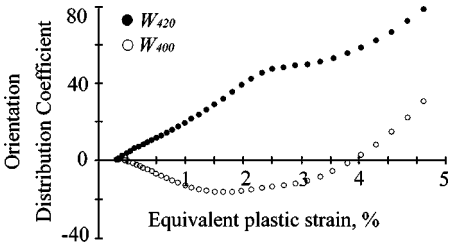


FIG. 5. Orientation distribution coefficient changes under uniaxial tension in the plastic deformation range.

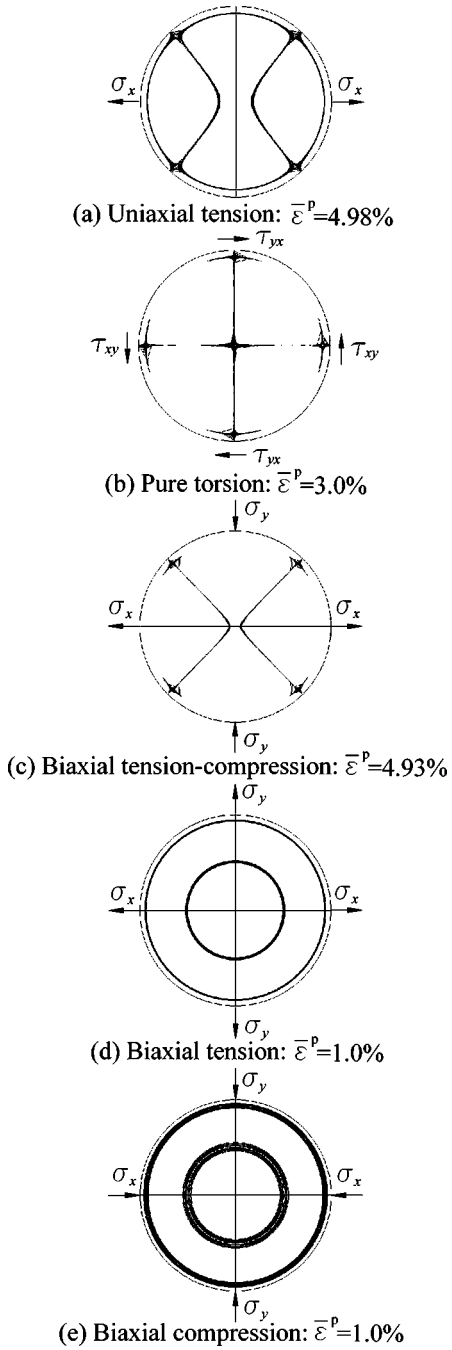


FIG. 6. {111} ultrasonic pole figures of an aluminum alloy based upon an ultrasonic nondestructive evaluation method.

observed at the outside area, except the 45° rotational discrepancy caused by the different loading conditions. This similarity is reflected by the similar transverse wave velocity changes among both states shown in Fig. 4(b).

Figure 6 indicates that texture developments under the plastic deformation can be theoretically analyzed by the simulated results of ultrasonic wave velocities. However, the ultrasonic pole figures shown in Fig. 6 fail to describe the fine structures. This is mainly caused by construction of the pole figures with only two lower-order ODCs W_{400} and W_{420}, as illustrated in Eq. (45). If higher-order ODCs such as W_{440} can be determined by using the results of Rayleigh surface wave velocity, it is capable of constructing ultrasonic pole figures with finer structures,[5,37] as what x-ray and neutron techniques can do. The successful trial on the prediction of texture development by ultrasonic pole figures implies the possibility of practical applications of the proposed theoretical modeling. For example, in aluminum can production, ultrasonic waves can be used as a process control monitor; namely, texture measurement after hot rolling. This appears quite promising.

Sayers's theory is based on the following assumptions: the velocities of ultrasonic waves traveling in the (macroscopic) orthotropic continuum are influenced by (1) single crystal moduli and (2) the preferred orientation of the crystallites. Other effects, such as grain boundaries, impurities, dislocations, and inhomogeneities are not taken into account in the modeling. Therefore, the accuracy and reliability of the theoretically analyzed results should be compared accordingly with the diffraction data of a neutron that penetrates through the metal samples as do the ultrasonics. However, the comparison with the neutron diffraction pole figure is not made in the present work, because the neutron equipments are usually available only at some central facilities. The comparison with x-ray diffraction data is also not made here, for the consideration of the x rays, sensing only a surface layer on the order of 10 μ thick. Therefore, an alternative method, i.e., a finite element polycrystal model (FEPM) is employed to examine the theoretically analyzed ultrasonic pole figures, because FEPM is already considered as a well-developed technology for the analysis of microstructural behaviors due to plastic deformation.

IV. COMPARISON OF ULTRASSONIC POLE FIGURES WITH POLE FIGURES ANALYZED BY FINITE ELEMENT POLYCRYSTAL MODEL (FEPM)

In this section, the finite element polycrystal model proposed by Takahashi *et al.*,[16] where each crystal is regarded as an element and its orientation is given randomly in the case of initial isotropy, is applied to calculate crystal orientations (Euler angles) of a face centered cubic (FCC) aggregate under plastic deformation. These calculated Euler angles (θ,ψ,ϕ) are then used to construct {111} pole figures via the stereographic projection method.

A. The stereographic projection method

In order to be able to discuss specific directions and planes within a crystal, it is very useful to have some sort of a map on which we can show these directions and planes of crystals. The stereographic projection provides us with such a map and it is widely used in metallurgical literature. Although much information on this subject can be found by several authors, for example by John,[38] it is still worthwhile to introduce briefly how stereographic projection is obtained.

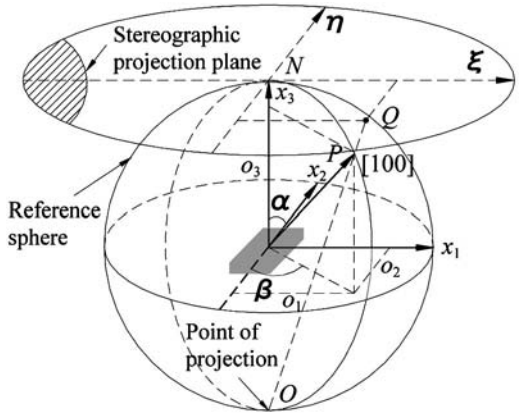


FIG. 7. Schematic diagram of stereographic projection method.

As it is shown in Fig. 7, a sample reference frame is defined first. A reference sphere of unit radius, centered at the origin of the sample reference frame, is then defined. The stereographic projection plane is defined as the tangent plane that contacts to the reference sphere at point N in the sample reference frame. In this system, a crystallographic direction, for example, [100], is represented by its pole P at which the direction intersects with the reference sphere. The pole P is, then, represented by its stereographic projection Q, which is the intersection of the projection plane with the line OP, where OP is the line from the point of projection O to the pole P.

As discussed in the preceding section, the microscopic coordinate X_i has a relationship with macroscopic coordinate x_i by Eqs. (38) or (39). However, it turns out to be more useful and convenient to specify the orientation of a crystal plane by Miller indices for stereographic projection. Now consider a crystal plane denoted by Miller indices (h_1,h_2,h_3). The cosine directions (O_1,O_2,O_3) of a unit normal vector **O** in this crystal plane can be expressed in terms of Miller indices as

$$O_i=\frac{h_i}{\sqrt{h_1^2+h_2^2+h_3^2}} \quad (i=1,2,3). \tag{49}$$

As shown in Fig. 7, the vector **o** with respect to the sample coordinate system is related to the unit normal vector **O** with respect to the crystallite coordinate system, according to Eq. (39), by

$$o_i=R_{ji}O_j \quad (i=1,2,3). \tag{50}$$

Let the orientation of **o** with respect to the sample coordinate system $o-x_1x_2x_3$ be specified by polar angle α and azimuthal angle β, as shown in Fig. 7; then its components can be obtained as

$$o_1=\sin\alpha\cos\beta, \quad o_2=\sin\alpha\sin\beta, \quad o_3=\cos\alpha. \tag{51}$$

Denote its stereographic projection as Q; we can easily get its coordinates as

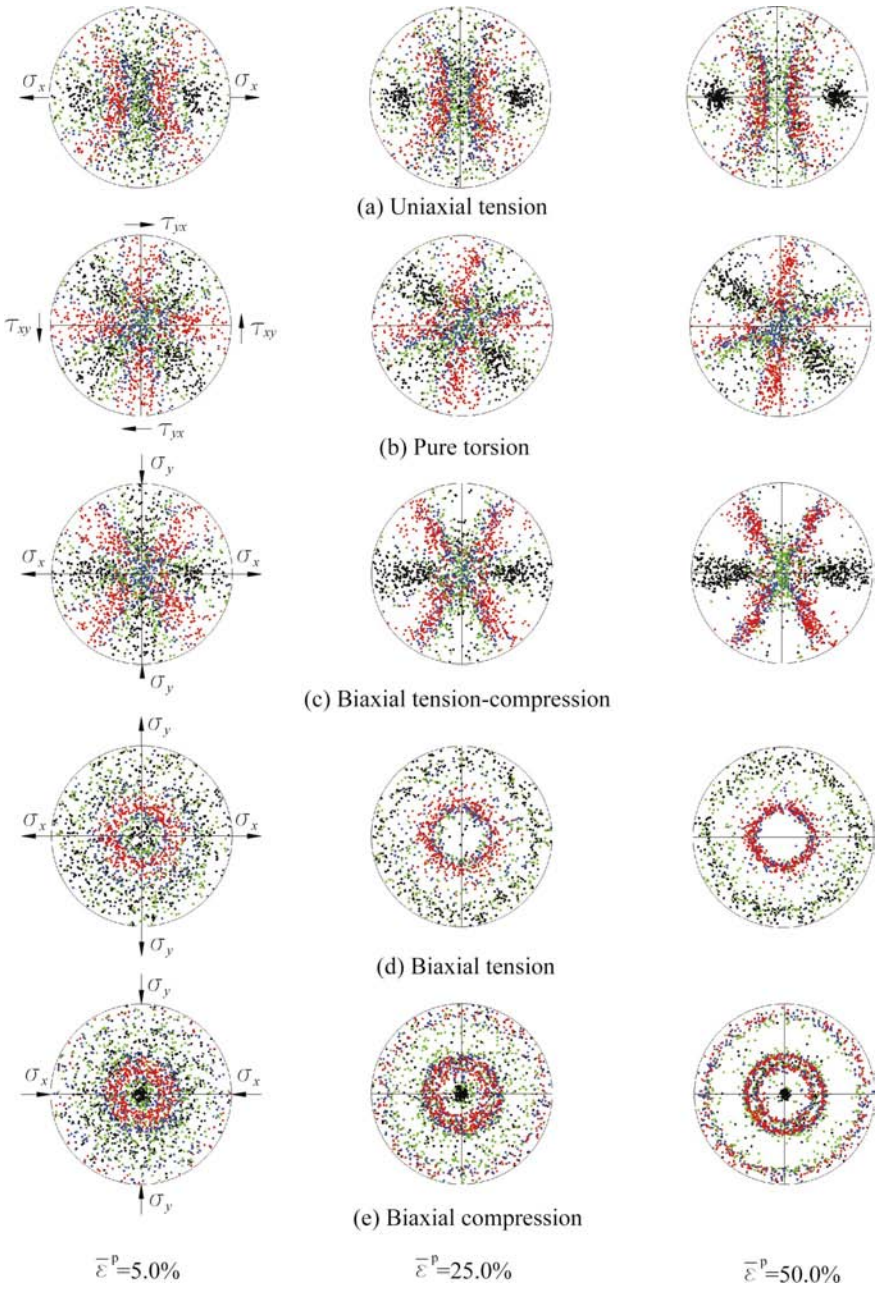


FIG. 8. {111} pole figures in FCC polycrystal analyzed by FEPM.

$$\xi = \frac{2\sin\alpha\cos\beta}{1+\cos\alpha} = \frac{2o_1}{1+o_3}, \quad \eta = \frac{2\sin\alpha\sin\beta}{1+\cos\alpha} = \frac{2o_2}{1+o_3}. \tag{52}$$

This indicates that if we know Euler angles (θ,ψ,ϕ), the unit normal vector **O** of any crystal plane can be represented by its stereographic projection.

B. {111} pole figures analyzed by finite element polycrystal model

Finite element analysis is carried out with $N=9\times9\times9=729$ elements possessing their random crystal orientations before deformation. With regard to the present purpose, Euler angle changes under various loading conditions of uniaxial tension, pure torsion, biaxial tension–compression, biaxial compression, and biaxial tension are calculated by FEPM. {111} pole figures are then constructed, in the way described above, and they are shown in Fig. 8.

Figure 8 shows the texture developments due to the plastic deformation (at the equivalent plastic strains 5%, 25%, and 50%) under various loading conditions. As it is well known that each crystal in FCC metals such as in aluminum possesses four slip planes, the slip system having a maximum slip amount among the four slip planes is called by the primary slip system; the slip systems on the other three planes are called by secondary slip systems. Hence, in pole figures as shown in Fig. 8, the primary slip system in each crystal is colored with red and three secondary slip systems are distinguished by colors of blue, green, and black according to the order of slip amount.

From Fig. 8 we can see that crystals having a randomly distributed orientation at an initial isotropic state show a preferred orientation as the development of plastic deformation

(25% and 50%). This implies that as the metal is deformed, the individual grains will rotate and slide with respect to each other under the applied forces, and as the development of plastic deformation, the polycrystalline aggregate will be changed from initial isotropy to strong anisotropy. Here it is also pointed out that the texture developments under pure torsion (a simple shear state) and biaxial tension–compression (pure shear state) are similar to each other, that is, both show a sharp tendency as the development of plastic deformation, as shown in Figs. 8(b) and 8(c). The discrepancy between the two pole figures lies in the fact that total spin $\dot{\omega}_{12}$ is assumed to be zero under a pure shear state, however, it is clearly different from zero under a simple shear state, as expressed by Eq. (25), reported by Takahashi *et al.*,[16] or other related articles by Kobayashi *et al.*[36] Except for this, the stress and strain states are the same among simple and pure shear states.

C. Comparison and discussion

By comparing Fig. 6 with Fig. 8, the good quantitative agreement between ultrasonic pole figures and FEPM pole figures of the primary system (denoted by a red color) is observed. This implies that the texture development of a primary slip system under plastic deformation can be theoretically predicted using W_{400} and W_{420}, determined by ultrasonic transverse wave velocities. However, the ultrasonic pole figures based upon only transverse waves presented in the present paper fail to describe the finer structures. This is mainly caused by consideration of only two lower-order ODCs to construct ultrasonic pole figures, as shown in Eq. (45). In fact, providing that higher-order ODCs such as W_{440} are determined from Rayleigh surface wave velocity, it is capable of constructing ultrasonic pole figures with a finer structure,[5,37] as what FEPM can do. However, regardless of such deficiencies, this research work suggested the possibility of applying our proposed theoretical modeling of an ultrasonic nondestructive evaluation method to the texture characterization, and also the possibility of clearing the physical meaning of the ultrasonic pole figure by comparing the ultrasonic pole figures with those analyzed by FEPM.

V. CONCLUSIONS

In the present paper, as one of authors' series study on an ultrasonic nondestructive evaluation method, we have investigated the frequency dependence of ultrasonic wave velocities via the Granato–Lücke model for dislocation damping. As a trial of its application, the proposed theoretical modeling was then applied to the texture characterization via ultrasonic pole figures under various loading conditions of uniaxial tension, pure torsion, biaxial tension–compression, biaxial compression, and biaxial tension. These predicted results were compared with pole figures analyzed by a finite element polycrystal model. Good quantitative agreement between the two results suggested the accuracy and reliability of our proposed theoretical modeling. As a result we can conclude the following.

(1) At small range of frequencies (less than 5 MHz), velocity changes of the longitudinal wave due to plastic deformation are independent on frequencies. The merit of this property enables us to study texture development by means of ultrasonic pole figures. On the other hand, velocity changes due to plastic deformation are almost negligible at high frequency; an appropriate construction of ultrasonic pole figures may be difficult via wave velocities at high frequencies.
(2) Texture changes caused by a primary slip system can be theoretically predicted by numerical simulations of transverse wave velocities. This implies the possibility of using the proposed theoretical modeling as a process control monitor in the manufacturing industry, for example, the texture measurement of a rolled alloy sheet. It appears quite promising.

ACKNOWLEDGMENTS

The research reported in this paper has been supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Culture, and Sport of Japan under Grant No. 11450039. The authors also acknowledge the helpful discussion with Professor Takahashi of Yamagata University and the generous assistance from Professor Ohashi, Associate Professor Fujiki, and Research Associate Miura of the Kitami Institute of Technology.

[1] R.-J. Roe, "Description of crystallite orientation in polycrystalline materials. III. General solution to pole figure inversion," J. Appl. Phys. **36**, 2024–2031 (1965).
[2] R.-J. Roe, "Inversion of pole figures for materials having cubic crystal symmetry," J. Appl. Phys. **37**, 2069–2072 (1966).
[3] H. J. Bunge, "Zur darstellung allgemeiner texturen (The general description on texture)," Z. Metallkd. **56**, 872–874 (1965).
[4] C. M. Sayers, "Ultrasonic velocities in anisotropic polycrystalline aggregate," J. Phys. D **15**, 2157–2167 (1982).
[5] M. Hirao, N. Hara, and H. Fukuoka, "Ultrasonic monitoring of texture in cold-rolled steel sheets," J. Acoust. Soc. Am. **84**, 667–672 (1988).
[6] A. V. Clark, Jr., R. C. Reno, R. B. Thompson, J. F. Smith, G. V. Blessing, R. J. Fields, P. P. Delsanto, and R. B. Mignogna, "Texture monitoring in aluminium alloys: a comparison of ultrasonic and neutron diffraction measurement," Ultrasonics **26**, 189–197 (1988).
[7] J. A. Szpunar and D. C. Hinz, "Correlating the plastic strain ratio with ultrasonic velocities in textured metals," J. Mater. Sci. **24**, 1233–1242 (1989).
[8] P. P. Delsanto, R. B. Mignogma, and A. V. Clark, "Ultrasonic texture and stress measurements anisotropic polycrystalline aggregates," J. Acoust. Soc. Am. **87**, 215–224 (1990).
[9] M. Kobayashi, "Acousto-plastic effects caused by elastic–plastic deformation of solids," Trans. Jpn. Soc. Mech. Eng., Ser. A **48**, 1072–1081 (1982) (in Japanese).
[10] M. Kobayashi, "Theoretical study of acousto-elastic effects caused by plastic anisotropy growth," Int. J. Plast. **3**, 1–20 (1987).
[11] M. Kobayashi, "Ultrasonic evaluation of microstructural changes of solid materials under plastic deformation," in *Characterization of Mechanical Properties of Materials*, edited by Wei-Yang Lu and Chi-Sing Man, ASME MD-Vol. 33, 33–46, 1992.
[12] M. Kobayashi, "Ultrasonic nondestructive evaluation of microstructural changes of solid materials under plastic deformation—Part I. Theory," Int. J. Plast. **14**, 511–522 (1998).
[13] M. Kobayashi, H. Murakami, and T. Kayaba, "Numerical and experimental studies on plastically induced acoustoelastic effects," in *Proceedings of the 6th Congress on Experimental Mechanics* (The Society for Experimental Mechanics, Portland, 1988), pp. 416–421.
[14] M. Kobayashi, "Ultrasonic nondestructive evaluation of microstructural changes of solid materials under plastic deformation—Part II. Experiment and simulation," Int. J. Plast. **14**, 523–535 (1998).
[15] A. Granato and K. Lücke, "Theory of mechanical damping due to dislocations," J. Appl. Phys. **27**, 583–593 (1956).

[16] H. Takahashi, H. Motohashi, M. Tokuda, and T. Abe, ''Elastic–plastic finite element polycrystal model,'' Int. J. Plast. **10**, 63–80 (1994).
[17] F. D. Murnaghan, *Finite Deformation of an Elastic Solid* (Wiley, New York, 1951).
[18] C. Truesdell and W. Noll, *The Nonlinear Field Theories of Mechanics*, 2nd ed. (Springer-Verlag, Berlin, 1992), p. 229.
[19] J. E. Marsden and T. J. R. Hughes, *Mathematical Foundation of Elasticity* (Prentice–Hall, Englewood Cliffs, NJ, 1983).
[20] A. A. Il'lushin, ''On the increments of plastic deformations and the yield surface,'' PMM **24**, 663–666 (1960).
[21] Y. F. Dafalias, ''Il'lushin's postulate and resulting thermodynamic conditions on elasto–plastic coupling,'' Int. J. Solids Struct. **13**, 239–251 (1977).
[22] G. Z. Voyiadjis, ''Degradation of elastic modulus in elastoplastic coupling with finite strains,'' Int. J. Plast. **4**, 335–353 (1988).
[23] E. H. Lee, ''Finite deformation effects in plasticity analysis,'' in *Plasticity Today, Modeling, Methods and Applications*, edited by A. Sawczuk and G. Bianchi (Elsevier Applied Science, Amsterdam, 1983), pp. 61–74.
[24] E. C. Aifantis, ''The physics of plastic deformation,'' Int. J. Plast. **3**, 211–247 (1987).
[25] S. Stören and J. R. Rice, ''Localized necking in thin sheets,'' J. Mech. Phys. Solids **23**, 421–441 (1975).
[26] J. Christoffersen and J. W. Hutchinson, ''A class of phenomenological corner theories of plasticity,'' J. Mech. Phys. Solids **27**, 465–487 (1979).
[27] M. Gotoh, ''A class of plastic constitutive equations with vertex effect, I. General theory,'' Int. J. Solids Struct. **21**, 1101–1116 (1985).
[28] M. Gotoh, ''A class of plastic constitutive equations with vertex effect, II. Discussions on the simple form,'' Int. J. Solids Struct. **21**, 1117–1129 (1985).
[29] M. Kuroda, ''Plastic spin associated with a corner theory of plasticity,'' Int. J. Plast. **11**, 547–570 (1995).
[30] H. P. Shrivastava, Z. Mroz, and R. Dubey, ''Yield criterion and the hardening rule for a plastic solid,'' Z. Angew. Math. Mech. **53**, 625–633 (1973).
[31] Y. Yoshimura, ''Hypothetical theory of anisotropy and the Bauschinger effects due to plastic strain history,'' Technical Report No. 349, Aero. Res. Inst. Univ. of Tokyo, Tokyo, 1959.
[32] M. Kobayashi, ''Ultrasonic nondestructive evaluation of annealed effects on plastic deformation of aluminum alloy,'' in *Proceedings of the 6th Conference on Mechanical Behaviour of Materials*, edited by M. Jono and T. Inoue (Pergamon, Oxford, 1991), pp. 735–740.
[33] M. Kobayashi, ''Acoustoelastic theory for plastically deformed solids,'' JSME Int. J. **33**, 310–318 (1990).
[34] J. F. Nye, *Physical Properties of Crystals: Their Representation by Tensors and Matrices* (Oxford University Press, Oxford, 1985), pp. 131–149.
[35] M. Kobayashi, S. Suzuki, S. Miura, and S. Oomori, ''Experimental verification of ultrasonic nondestructive material evaluation method under combined stress states and evaluation of plastic anisotropic growth,'' Trans. Jpn. Soc. Mech. Eng., Ser. A **65**, 2149–2156 (1999) (in Japanese).
[36] M. Kobayashi, S. Tang, S. Miura, K. Iwabuchi, S. Oomori, and H. Fujiki, ''Ultrasonic nondestructive material evaluation method and study on texture and cross slip effects under simple and pure shear states,'' Int. J. Plast. **19**, 771–804 (2003).
[37] H. Fukuoka, H. Toda, and M. Hirao, *Foundations and Applications of Acoustoelasticity* (Ohmsha, Tokyo, 1993), pp. 107–150 (in Japanese).
[38] D. V. John, *Fundamentals of Physical Metallurgy* (Wiley, New York, 1975), pp. 5–28.

The effects of membrane metallization in capacitive microfabricated ultrasonic transducers

Alessandro Caronti,[a] Riccardo Carotenuto, Giosuè Caliano, and Massimo Pappalardo
Dipartimento di Ingegneria Elettronica, Università Roma Tre, Via della Vasca Navale 84, 00146 Roma, Italy

(Received 12 January 2003; revised 15 November 2003; accepted 24 November 2003)

The mechanical effects of the metal layer on the membranes of capacitive micromachined ultrasonic transducers (CMUTs) are analyzed in this paper by means of finite element simulations. The influence of electrode size and thickness on the electrostatic behavior of the single CMUT cell, including diaphragm displacement, cell capacitance, and collapse voltage, is explored. The effect on device sensitivity is investigated through the transformation factor of the cell, that is computed by FEM and compared with the parallel plate model prediction. It is found that for a non-negligible electrode thickness, as in the majority of fabricated devices, both the static and dynamic performance of the cell can be affected in a significant way. Thus, the effects of membrane metallization must be taken into account in CMUT design and optimization. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1642622]

PACS numbers: 43.38.Bs, 43.38.Ar [AJZ] Pages: 651–657

I. INTRODUCTION

Capacitive ultrasonic transducers have been used for decades for the generation and detection of ultrasonic waves. In recent years, advances in microfabrication techniques have led to a new generation of capacitive micromachined ultrasonic transducers (CMUTs),[1] that are potentially competitive with the piezoelectric transducers. The main advantages of CMUTs compared to piezoelectrics are the increased bandwidth in immersion, the integration with electronics, and the ability to fabricate large two-dimensional (2-D) arrays for real-time volumetric imaging. However, many applications, especially in medical imaging, demand a better understanding and improvements of the CMUT design. This paper is focused on the mechanical effects of the metal layer on the membranes of CMUTs.

There are relatively few publications concerning the electrode design in capacitive transducers.[2–4] Voorthuyzen *et al.*[2] have optimized the electrode size of capacitive pressure sensors with respect to sensitivity and signal-to-noise ratio. Bozkurt *et al.*[3] have presented the optimization criteria for CMUTs as a trade-off between sensitivity and bandwidth requirements, leading to an optimum design of the electrode radius ranging from 40% to 50% of the membrane radius. More recently, Fletcher and Thwaites[4] have proposed a new design of condenser microphones in which the bottom electrode has a curved profile instead of planar, resulting in increased sensitivity and low harmonic distortion. However, the mechanical load of the top electrode on the diaphragm was not considered in previous works.

Indeed, because the electrode thickness cannot be made negligible in comparison with the membrane thickness in CMUTs, both the static and dynamic performance of the single CMUT cell are found to be considerably affected by the metal layer. The effects of membrane metallization on the static displacement of the diaphragm, its mechanical resonant frequency, the cell capacitance, and the collapse voltage, are explored in this paper by means of a finite element analysis. The influence of the electrode size on device sensitivity is investigated through the transformation factor of the transducer. A general definition of the transformation factor is presented, along with an analytical derivation for an electrostatic cell with a metallized circular diaphragm; an accurate computation is then performed through finite element simulations.

The results of the analysis demonstrate that the mechanical effects of the metal layer must be taken into account in CMUT design and optimization.

II. THE TRANSFORMATION FACTOR OF A CAPACITIVE TRANSDUCER

One of the most important parameters affecting the performance of a capacitive transducer is the transformation factor ϕ, which can be defined, in a general way, as the slope of the force–voltage curve:

$$\phi = \left| \frac{\partial F_E}{\partial V} \right|_{V=V_{\mathrm{dc}}}, \tag{1}$$

where F_E is the dc electrostatic force arising from the bias voltage V_{dc}. This parameter appears in the electromechanical equivalent circuit of Fig. 1 as the turns ratio of the ideal transformer, thus relating the amplitude of the driving electrostatic force on the mechanical side to the amplitude of the alternating voltage on the electrical side. Both the transmitting and receiving voltage sensitivities are proportional to the transformation factor within the validity of a linear model.[5,6]

A CMUT electrostatic cell basically consists of a metallized silicon nitride membrane placed a very small distance apart from a silicon substrate; many such cells are electrically connected in parallel to make a transducer element, as

[a]Telephone: +39 06 55177081; fax: +39 06 5579078; electronic mail: caronti@uniroma3.it

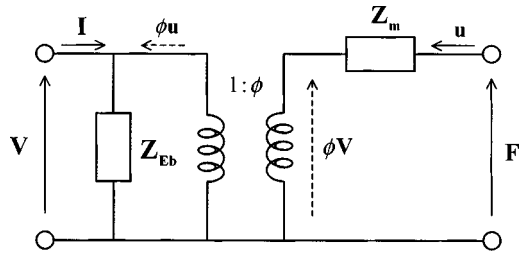


FIG. 1. Electromechanical equivalent circuit of a capacitive transducer. $\mathbf{Z_{Eb}}$ is the blocked electrical impedance, $\mathbf{Z_m}$ is the mechanical impedance. According to phasor notation, boldface is used to represent quantities in the frequency domain.

shown in Fig. 2(a), and the elements are used to make CMUT arrays [Fig. 2(b)]. If the single cell is represented as a parallel plate capacitor and the fringing field is neglected, the electrostatic force is given by

$$F_E = -\frac{\epsilon_0 A V^2}{2 d_e^2}, \tag{2}$$

where ϵ_0 is the permittivity of vacuum, A is the area of the plates, and d_e is the equivalent distance between the electrodes, accounting for the various dielectric materials placed between them. Equation (2) assumes a fully metallized membrane and a uniform electrostatic pressure over its surface.

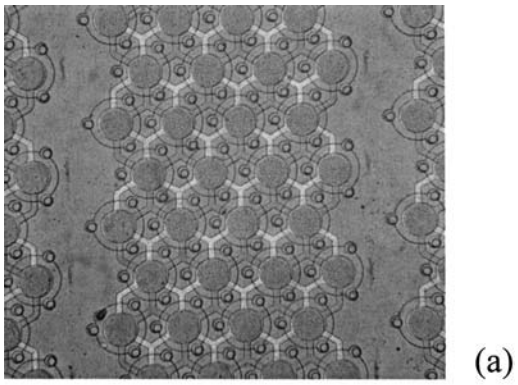


FIG. 2. The top view of a one-dimensional (1-D) CMUT array. (a) Magnified view of an element, consisting of a 2-D array of circular membranes. The regions in dark gray are membrane top electrodes and interconnections, the regions in light gray are the bottom metallizations, the holes around each membrane are used to form the cavity below by selective etching of a sacrificial material. (b) A portion of a 64 elements CMUT array (Ref. 9).

By differentiating Eq. (2) with respect to the voltage V, the classical expression of ϕ can be obtained,[5]

$$\phi = \frac{C_0 V_{\mathrm{dc}}}{d_e} = \frac{C_0^2 V_{\mathrm{dc}}}{\epsilon_0 A}, \tag{3}$$

where $C_0 = \epsilon_0 A / d_e$ is the capacitance of the polarized cell.

Actually, the electrostatic force deflects the membrane and the effective distance d_e is a function of the position. Moreover, the membrane can be partially metallized, so that Eq. (2) can be rewritten for a circular geometry as

$$\begin{aligned} F_E &= -\iint_{S_m} \frac{\epsilon_0 V^2}{2[d_e - w(r)]^2} dS_m \\ &= -\pi \epsilon_0 \int_0^{a_{\mathrm{el}}} \frac{V^2}{[d_e - w(r)]^2} r\, dr, \end{aligned} \tag{4}$$

where S_m is the metallized area, a_{el} is the electrode radius, and $w(r)$ is the deformed shape of the membrane resulting from the applied voltage V. The fringing field is still neglected in Eq. (4).

Substituting Eq. (4) into Eq. (1), we obtain for the transformation factor of a cell with a circular membrane the following analytical expression:

$$\begin{aligned} \phi &= \pi \epsilon_0 \left| \int_0^{a_{\mathrm{el}}} \frac{\partial}{\partial V} \left\{ \frac{V^2}{[d_e - w(r)]^2} \right\} r\, dr \right|_{V = V_{\mathrm{dc}}} \\ &= 2\pi \epsilon_0 \left| \int_0^{a_{\mathrm{el}}} \frac{V}{[d_e - w(r)]^2} \right. \\ &\quad \left. \times \left(1 + \frac{V}{[d_e - w(r)]} \frac{\partial w(r)}{\partial V} \right) r\, dr \right|_{V = V_{\mathrm{dc}}}, \end{aligned} \tag{5}$$

which can be rewritten in terms of the transverse electric displacement in the gap, D_z, as follows:

$$\phi = 2\pi \left| \int_0^{a_{\mathrm{el}}} \frac{D_z(r)}{[d_e - w(r)]} \left(1 + \frac{D_z(r)}{\epsilon_0} \frac{\partial w(r)}{\partial V} \right) r\, dr \right|_{V = V_{\mathrm{dc}}}. \tag{6}$$

As can be seen, the transformation factor depends on membrane deflection and electric field distribution in the gap, as well as on the electrode radius. The electrode thickness t_{el} is also an important variable that affects both the electrical impedance of the transducer and the electrostatic behavior of the cell. The value of t_{el} is chosen as a compromise to reduce the electrical resistance of the interconnections among the CMUT cells without loading the membranes in a significant way; t_{el} ranges between 0.1 and 0.3 μm in fabricated devices.[7–9]

In the next section, a finite element calculation of the transformation factor is presented that quantitatively demonstrates the possibility to enhance the performance of CMUTs by electrode patterning.

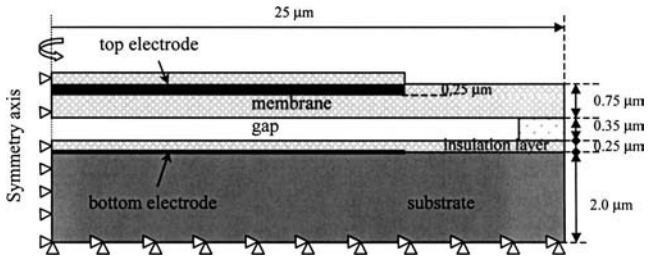


FIG. 3. The FEM model of a single CMUT cell.

III. FEM SIMULATION RESULTS

The finite element analysis of the electrostatic behavior of a single CMUT cell was performed using the commercial package ANSYS 6.1 (ANSYS Inc., Canonsburg, PA). The FEM axisymmetric model of the cell, shown in Fig. 3, includes a silicon nitride membrane of 25 μm radius and 0.75 μm thickness with an aluminum top electrode of 0.25 μm thickness, a thin (0.25 μm) silicon nitride insulation layer over the bottom electrode, and a 0.35 μm vacuum gap between the membrane and the insulation layer. The bottom electrode is assumed to be infinitesimally thick, because it does not affect the membrane displacement and the electrostatic solution. The membrane is supposed without tensile stress, and it is supported on a silicon oxide post along its rim; the silicon substrate (assumed thin for simplicity) is clamped at the bottom.

The mechanical and electrical material properties used in the simulations are listed in Table I. The permittivity was the only input parameter in the electrostatic analysis that was performed using the two-dimensional electrostatic elements PLANE121; the eight-node solid elements PLANE82 were used in the structural analysis.

The problem of coupled electrostatic–mechanical fields involved in capacitive transducer simulations was solved using the ANSYS macro ESSOLV. This solver performs a coupled electrostatic–structural analysis by automatically iterating between an electrostatic field solution and a structural solution until the field and the structure are in equilibrium. The solution converges when both the maximum structural displacement and stored electrostatic energy change by less than a specified tolerance level (e.g., 0.5%) between the current and previous iteration.

A. Effect of electrode size and thickness on membrane deflection

Figure 4 compares the membrane displacement profiles produced by a polarizing voltage of 80 V for three electrode sizes, assuming a zero thickness electrode (top) and a 0.25 μm thickness (bottom). The deformed shapes are plotted as a function of the radial position, ranging from the center ($r=0$) to the edge of the membrane ($r=25\,\mu$m). In the case of a zero thickness electrode [see Fig. 4(a)], the membrane deflection increases progressively with the electrode radius, as expected; on the contrary, with a 0.25 μm metal thickness [see Fig. 4(b)], the displacement achieved with an electrode radius of 18 μm is higher than that of a fully metallized membrane. Thus, for a non-negligible metal thickness, a partial membrane metallization can produce larger deformations than a full metallization. The reason is that, as the electrode radius increases, the resulting increase of the electrostatic force is opposed by the stiffening of the diaphragm, which can bend to a smaller extent. Figure 5 shows a comparison between the center displacement values for the two cases of zero metal thickness and 0.25 μm thickness, pointing out the effect of electrode size on the mechanical compliance of the SiN–Al composite diaphragm.

TABLE I. Electrical and mechanical material properties used in the finite element simulations.

Parameters	SiN	SiO	Si	Aluminum
Young's modulus (GPa)	280	310	169	67.6
Poisson's ratio	0.26	0.2	0.3	0.35
Density (kg/m^3)	3200	2220	2330	2700
Relative permittivity	7.0	4.0	11.8	

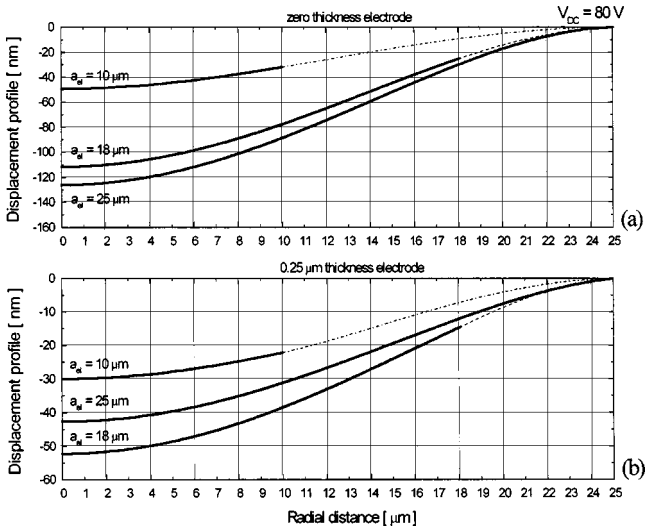


FIG. 4. Deformed membrane shapes with (a) zero thickness electrode and (b) the 0.25 μm thickness electrode.

B. Transformation factor, collapse voltage, and cell capacitance

The transformation factor ϕ of the cell was calculated in three different ways, according to Eqs. (1), (3), and (6). The

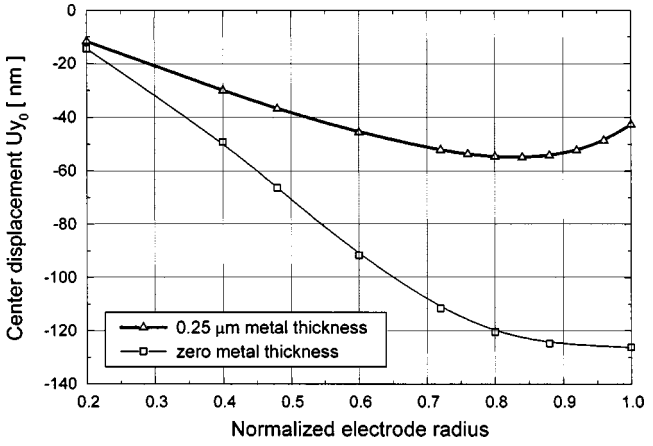


FIG. 5. Center displacement of the diaphragm produced by an 80 V bias voltage as a function of the electrode radius, normalized with respect to the membrane radius.

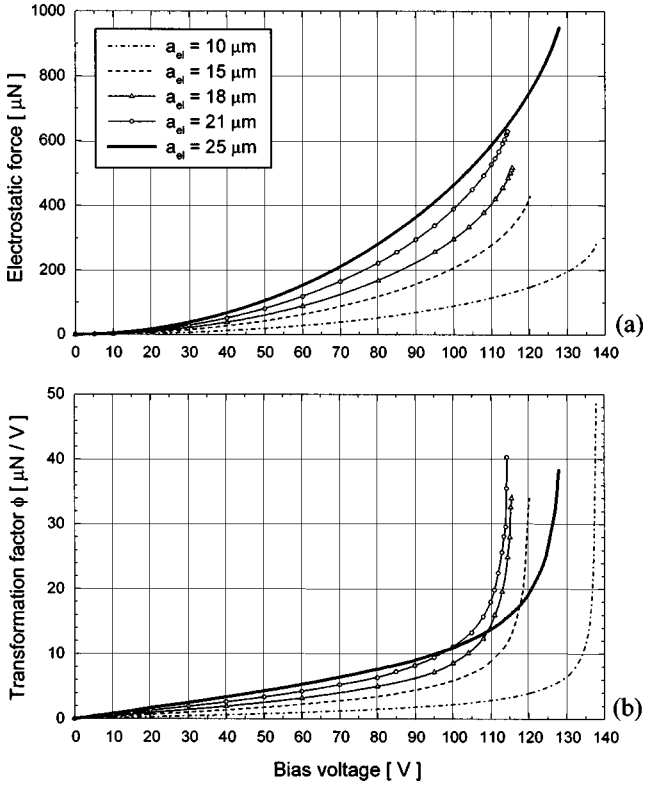


FIG. 6. (a) Electrostatic force and (b) transformation factor of the modeled CMUT cell for different sizes of the electrode.

electrostatic force F_E of Eq. (1) was computed at the gap–membrane interface as the product of the induced surface charge density and the electric field. In general, in a Cartesian coordinate system, the component of electrostatic force parallel to the i axis is given by

$$F_{Ei}=\int_{\Gamma} T_{ij}^{\mathrm{M}} n_j \, d\Gamma, \quad i,j=1,2,3, \tag{7}$$

where Γ is the integration surface (the interface), n_j is the jth component of the unit vector normal to the surface element $d\Gamma$, and T_{ij}^{M} is the Maxwell stress tensor, defined as[10]

$$T_{ij}^{\mathrm{M}}=E_i D_j-\frac{1}{2}\,\delta_{ij} E_k D_k, \quad i,j,k=1,2,3. \tag{8}$$

Here, E_i and D_j denote the components of the electric field and electric displacement, respectively; δ_{ij} denotes the Kronecker delta: $\delta_{ij}=1$ when $i=j$ and zero otherwise. In Eqs. (7) and (8) the summation convention is adopted, i.e., a repeated index implies summation over its range.

For the modeled structure and material assumptions, the electric displacement vector is always parallel to the electric field, and the direction of the unit vector n is nearly constant over the membrane surface for small deflections. Therefore, within the accuracy of the FEM computation, the transverse electrostatic force can be approximated to the component perpendicular to the equilibrium plane of the membrane. A plot of the normal component of the electrostatic force as a function of the bias voltage for increasing electrode sizes is shown in Fig. 6(a); according to Eq. (1), the voltage derivative of these curves yields the transformation factor curves, as reported in Fig. 6(b).

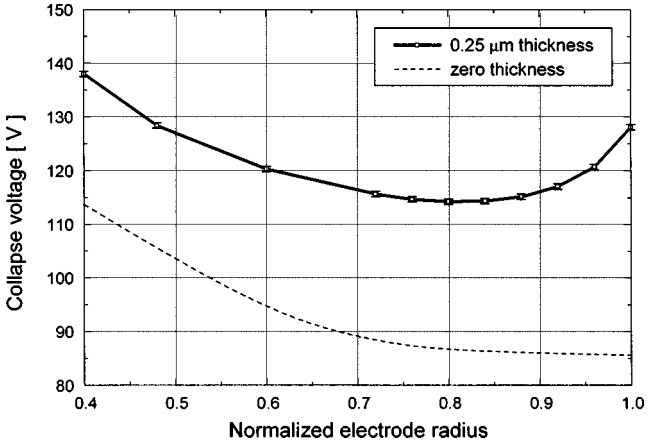


FIG. 7. Collapse voltage as a function of the normalized electrode radius.

The maximum value of the dc voltage, for each electrode size, is the voltage at which the membrane collapses onto the substrate, which is usually called collapse (or critical) voltage;[11,12] when the bias voltage is higher than the collapse voltage, the successive iterations of the macro ESSOLV do not converge to a stable deformed shape. A plot of the collapse voltage for the cell of Fig. 3 is shown in Fig. 7 as a function of the electrode radius, normalized with respect to the membrane radius. The collapse voltage was computed with an uncertainty of ± 0.5 V, as shown by the error bars in the figure. When the electrode size increases, the collapse voltage decreases up to a metallized area of about 80%; for larger metallization patterns, the stiffening effect of the composite diaphragm becomes important and the collapse voltage increases. However, both a half-metallized and a totally metallized membrane have a collapse voltage only 12% greater than its minimum value of about 114 V. It must be emphasized that these results depend significantly on the electrode thickness; for a zero metal thickness, the collapse voltage would decrease progressively with increasing the electrode radius, as indicated by the dashed line in Fig. 7. Also note that the behavior of the collapse voltage as a function of the electrode size follows that of the center displacement of the diaphragm, as shown in Fig. 5: the higher the deflection achieved with a given bias voltage, the lower the voltage that brings the diaphragm into collapse.

As a further validation of the theory and the simulation results presented, the transformation factor was also computed by means of Eq. (6), evaluating through the FEM model the electric displacement in the gap, D_z, and the voltage derivative of the membrane displacement, $\partial w(r)/\partial V$. Figure 8 compares different simulations of ϕ, for an electrode radius of 22 μm, obtained according to Eqs. (1), (3), and (6). An excellent agreement exists between the direct calculation of Eq. (1) and that of Eq. (6); on the contrary, the simple relation given by Eq. (3), based on the parallel plate approximation, is not adequate when the bias voltage increases close to the collapse voltage. The discrepancy mainly arises from the term of Eq. (6), including the rate of increase of the displacement with the applied voltage, $\partial w(r)/\partial V$, that is not taken into account by Eq. (3).

The cell capacitance C_0 of Eq. (3) was computed for

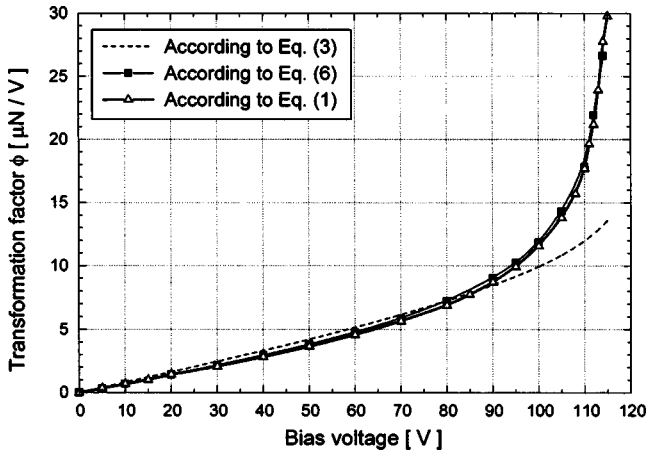


FIG. 8. A comparison of simulations of the transformation factor ϕ according to Eqs. (1), (3), and (6).

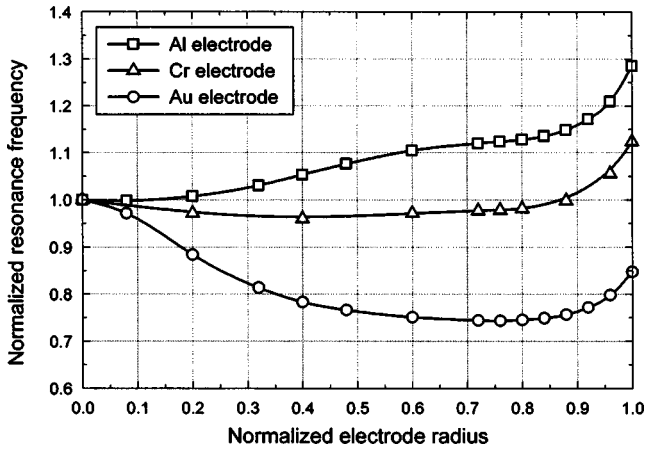


FIG. 10. Effects of metal type and size on the fundamental resonance frequency.

different values of the bias voltage using the ANSYS macro CMATRIX, relating the charges on the electrodes with the voltage drop. A plot of the capacitance versus the polarizing voltage is shown in Fig. 9 for several values of the electrode radius. It should be noted that the contribution of the interconnections between adjacent cells, as it is seen in Fig. 2(a), is not taken into account by the FEM model of Fig. 3. However, this contribution is negligible in the CMUT configuration of Fig. 2(a), because the top and bottom interconnections do not overlap.

Figure 9 shows that reducing the metal covering from 100% ($a_{el}=25\,\mu$m) to 40% ($a_{el}=10\,\mu$m), the static capacitance decreases more than five times, from 0.044 to 0.008 pF. Further, as the bias voltage is increased up to the collapse voltage, the initial value of the capacitance increases of 15%–30%, depending on the electrode radius.

C. Influence of the electrode on the mechanical resonance frequency

The fundamental resonance frequency of the composite diaphragm is influenced by the metal type and the electrode size in the way shown in Fig. 10. The resonance frequency is normalized with respect to that of a non-metallized diaphragm (or a diaphragm with a zero thickness electrode),

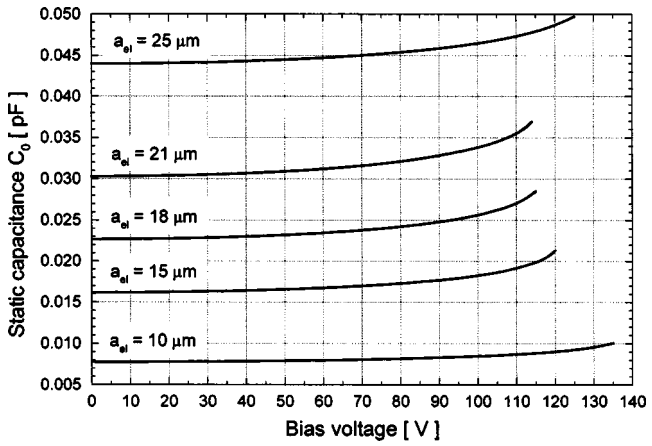


FIG. 9. Static capacitance versus the polarizing voltage for several electrode sizes.

whose value is calculated as 5.26 MHz. Three metal types are considered in the simulations: aluminum ($\rho_{Al}=2700\,\text{kg/m}^3$, $Y_{Al}=68\,\text{GPa}$), chromium ($\rho_{Cr}=7200\,\text{kg/m}^3$, $Y_{Cr}=140\,\text{GPa}$), and gold ($\rho_{Au}=19700\,\text{kg/m}^3$, $Y_{Au}=80\,\text{GPa}$).

As the electrode radius increases, the mass loading of the top electrode can affect significantly the resonance frequency, depending on the metal density and rigidity. In the case of a gold electrode, the resonance frequency is lowered down to 74% of its unloaded value, corresponding to a metallized area of about 75%; for larger electrode sizes, the mass loading effect is partly compensated by the stiffening effect of the composite diaphragm, and the resonant frequency slightly increases. On the other hand, for the aluminum electrode, whose density is much lower and Young's modulus very close to that of gold, the stiffening effect is dominant and the resonance frequency gradually increases with the electrode size, with a maximum deviation of 30%. The behavior of the SiN–Cr diaphragm is intermediate between these two cases, and the influence of the electrode size on the resonance frequency is negligible up to 90% of metal covering. Note that only the dynamic response of the diaphragm is affected by the metal density, not the electrostatic solution: in fact, the static displacement is determined by the elastic restoring force, which is only controlled by the stiffness of the composite plate.

D. Effects of geometric nonlinearities and in-plane stress

Geometric nonlinearities may arise in the structure due to the changing shape of the diaphragm as it deflects. The PLANE82 solid elements support the large displacement capability, and a simulation was run for an estimation of this effect. The center displacement of the diaphragm with an 80% metallization is shown in Fig. 11(a) as a function of the bias voltage, including or not the geometric nonlinearities. As it is seen, the resulting effect is a reduction of the collapse voltage by less than 2%, so that the transformation factor curve simply shifts to the left [see Fig. 11(b)]. This small variation is not unexpected, because the theoretical maximum center deflection before membrane collapse is only

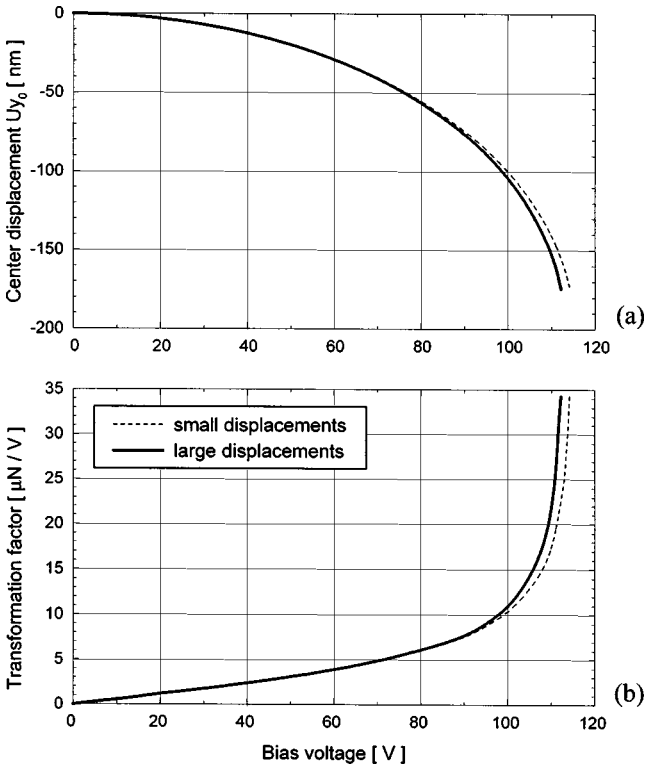


FIG. 11. (a) Center displacement of the diaphragm and (b) transformation factor of the cell, including large deflection effects. The broken curve shows the solution calculated assuming small deflections, while the full curve takes account of nonlinearities caused by changes in diaphragm curvature.

one-third of the effective gap height, in the parallel plate approximation.[11,12] Therefore, for the typical dimensions of a CMUT cell with a gap in the submicron range, the maximum center deflection is small as compared with the thickness of the plate (usually higher than 1 μm), and the large strain effects can be neglected.[13]

It should be noted that the results reported were obtained assuming the composite diaphragm with zero tensile stress, so that its elastic restoring force is only controlled by the structural stiffness of the materials. Actually, the deposition of thin silicon nitride films to form the CMUT membranes gives rise to stresses, which can be either tensile or compressive, depending on the parameters of the deposition process. Both in-plane stress and diaphragm thickness affect the resonant frequency and the voltage at which electrostatic pull-in occurs. To have a more robust control over the cell characteristics and transducer performance, a low-tensile stress is usually preferable.

Several techniques to control and reduce the stress in silicon nitride films can be employed: in Ref. 8, a post-deposition thermal annealing was used in a plasma-enhanced chemical vapor deposition (PECVD) process to change the stress from compressive to slightly tensile; in Ref. 14, the silane and ammonia ratio in the gas flow composition of a low-pressure chemical vapor deposition (LPCVD) was controlled such that a low residual stress was achieved.

Residual stress has also an important effect on the membrane shape after release, resulting in a deflection at the center due to the bending of the posts supporting the membrane at the edge. The CMUT membranes can bend upward or downward, depending on whether the stress is compressive or tensile.[8] As a consequence, the collapse voltage can either increase or decrease and, similarly to the behavior shown in Fig. 11(b), the transformation factor curve of the cell shifts right or left.

IV. CONCLUSIONS

A finite element study of the effects of membrane metallization in capacitive micromachined ultrasonic transducers is reported. The metal type, size, and thickness can affect significantly the performance of the single electrostatic cell, both in static and dynamic conditions. The importance of the analysis originates from the fact that, for practical reasons, the electrode thickness cannot be made negligible compared to that of the membrane. For a typical configuration of a CMUT cell, it is found that the collapse voltage and the displacement of the diaphragm change with the electrode size, depending on its thickness. The mass loading of the top electrode also affects the mechanical resonance frequency of the diaphragm, in a way that depends on the electrode dimensions and metal type. Further, the influence of the electrode on device sensitivity is investigated through the transformation factor, which is accurately computed by FEM. A comparison with the result of the parallel plate model shows a significant deviation as the bias voltage approaches the collapse voltage.

In summary, the effects of membrane metallization must be taken into account in CMUT design. The optimization of electrode size and shape to improve the performance of CMUTs will be the object of a future paper.

[1] M. I. Haller and B. T. Khuri-Yakub, ''A surface micromachined ultrasonic air transducer,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1–6 (1996).
[2] J. A. Voorthuyzen, A. J. Sprenkels, A. G. H. van der Donk, P. R. Scheeper, and P. Bergveld, ''Optimization of capacitive microphone and pressure sensor performance by capacitor–electrode shaping,'' Sens. Actuators A **25**, 331–336 (1991).
[3] A. Bozkurt, I. Ladabaum, A. Atalar, and B. T. Khuri-Yakub, ''Theory and analysis of electrode size optimization for capacitive microfabricated ultrasonic transducers,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **46**, 1364–1374 (1999).
[4] N. H. Fletcher and S. Thwaites, ''Electrode surface profile and the performance of condenser microphones,'' J. Acoust. Soc. Am. **112**, 2779–2785 (2002).
[5] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982).
[6] A. Caronti, G. Caliano, A. Iula, and M. Pappalardo, ''An accurate model for capacitive micromachined ultrasonic transducers,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **49**, 159–168 (2002).
[7] X. Jin, I. Ladabaum, F. L. Degertekin, S. Calmes, and B. T. Khuri-Yakub, ''Fabrication and characterization of surface micromachined capacitive ultrasonic immersion transducers,'' IEEE/ASME J. Microelectromech. Syst. **8**, 100–114 (1999).
[8] E. Cianci, V. Foglietti, D. Memmi, G. Caliano, A. Caronti, and M. Pappalardo, ''Fabrication of capacitive ultrasonic transducers by a low temperature and fully surface-micromachined process,'' J. Int. Soc. Precision Eng. Nanotechnol. **26**, 347–354 (2002).
[9] G. Caliano, R. Carotenuto, A. Caronti, M. Pappalardo, V. Foglietti, E. Cianci, L. Visigalli, and I. Persi, ''CMUT echographic probes: design and fabrication process,'' *IEEE Ultrasonics Symposium* (IEEE, New York, 2002), pp. 1040–1043.
[10] A. Nathan and H. Baltes, *Microtransducer CAD. Physical and Computational Aspects* (Springer-Verlag/Wien, New York, 1999).

[11] F. V. Hunt, *Electroacoustics. The Analysis of Transduction, and Its Historical Background* (American Institute of Physics, Woodbury, NY, 1982).
[12] A. Caronti, R. Carotenuto, and M. Pappalardo, ''Electromechanical coupling factor of capacitive micromachined ultrasonic transducers,'' J. Acoust. Soc. Am. **113**, 279 (2003).
[13] S. P. Timoshenko and S. W. Woinowski-Krieger, *Theory of Plates and Shells* (McGraw-Hill, New York, 1959).
[14] X. Jin, I. Ladabaum, and B. T. Khuri-Yakub, ''The microfabrication of capacitive ultrasonic transducers,'' IEEE/ASME J. Microelectromech. Syst. **7**, 295–302 (1998).

A tri-modal directional transducer[a)]

John L. Butler and Alexander L. Butler[b)]
Image Acoustics, Inc., Cohasset, Massachusetts 02025

Joseph A. Rice
Space and Naval Warfare Systems Center, San Diego, California 92152 and Naval Postgraduate School, Monterey, California 93943

(Received 8 May 2003; revised 29 October 2003; accepted 6 November 2003)

Acoustic transducers made from piezoelectric ceramic cylinders usually exploit the breathing or omnidirectional (omni) mode of vibration. However, with suitable voltage distribution, higher order extensional modes of the cylinder can be excited which produce directional radiation patterns. These modal radiation patterns can then be combined to synthesize desired beam patterns which may be steered by incrementing the excitation. This paper describes a model for the combined acoustic response of the extensional modes of vibration of a piezoelectric ceramic cylinder, a method of synthesizing a desired radiation pattern, and an experimental implementation of a directional transducer that uses these techniques. This tri-modal transducer is broadband and directional with a frequency independent beam pattern yet simple, small, and lightweight. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639326]

PACS numbers: 43.38.Fx, 43.40.Rj, 43.30.Yj [AJZ] Pages: 658–665

I. INTRODUCTION

Directional transmit or receive beam patterns are often desired in transducer applications. A common method is to use a large planar array of transducers where the pressure adds in-phase in the broadside direction but, because of phase differences, only partially adds in other directions leading to a directional beam. The beam may be steered by mechanically rotating the array or electrically introducing phase shifts on the transducer elements of the array. Another method is to use a large curved array with narrow beam steering through incremental activation of groups of piston transducer elements of the array. A broad beam variation replaces the piston transducer elements with circular rings laid out in the form of a cylindrical array with steering through separate excitation of individual rings.[1] An alternative method which avoids the need for a large array of separate elements is to construct a directional beam pattern through the addition of higher order modes of vibration from a single transducer. This method has been used in underwater acoustics to create orthogonal dipole mode and omni mode beam patterns from piezoelectric hollow cylinders.[2] It has also been shown[3] that highly directional optimized beams can, in principle, be obtained from piezoelectric spheres.[4] In this paper we present a simple means for creating directional beams from a cylinder using the quadrupole mode as well as the omni and dipole modes. The resulting transducer may be steered by changing the voltage amplitude rather than the phase on the electrodes of the cylinder.

Radially poled piezoelectric cylinders with continuous electrodes on the inner and outer surfaces are widely used in underwater acoustic transducers as single mode devices.[5] As projectors, cylinder transducers are operated typically in the vicinity of their first breathing mode resonance, where the surface velocity is uniform and the sound field generated is omni-directional in the plane perpendicular to the axis. Although usually operated below resonance, most cylindrical hydrophones also use this omni-directional mode. Ehrlich and Frelich[2] describe a cylindrical sensor that has a pattern of four electrodes, each covering a quadrant of the cylinder, that, when combined, generate orthogonal directional dipoles and omni receive beam patterns. Spherical sensors with patterned electrodes[4] as well as dual mode flextensional transducers[6–8] have also been used to achieve directionality through the omni and dipole modes. Although higher order modes have been used for super-directive microphones[9] and considered for PVDF hydrophones,[10] the implementation of higher modes in cylindrical projectors appears to have been overlooked.

The dynamics of the multimode cylinder can be represented by the two-dimensional planar motion of a ring. This is a reasonable approximation for a cylinder at resonance that is short and thin walled. The ''breathing'' mode of a ring is the fundamental extensional mode which occurs when extensional waves in the structure of the ring complete one wavelength around the circumference. The extensional strain is then in phase at all circumferential locations and this uniform extensional strain is translated, by the geometry, into a purely radial displacement of the ring wall. The resonance frequency of this fundamental mode f_0 is given by $c/\pi D$, where c is the speed of sound for extensional waves in the ring wall and D is the mean diameter.

The higher order extensional modes of a ring[11] have a radial displacement that is harmonic in the azimuth angle φ with an even component proportional to $\cos n\varphi$ and an odd component proportional to $\sin n\varphi$, where n is the mode number. The fundamental ''breathing'' mode has a mode number $n=0$. The second mode, $n=1$, has two nodes with the radial

[a)]Portions of this article were presented at the 143rd meeting of the Acoustical Society of America, Pittsburgh, Pennsylvania, June 6, 2002 [J. Acoust. Soc. Am. **111**, 2444 (2002)].
[b)]Electronic mail: jbutler@imageacoustics.com

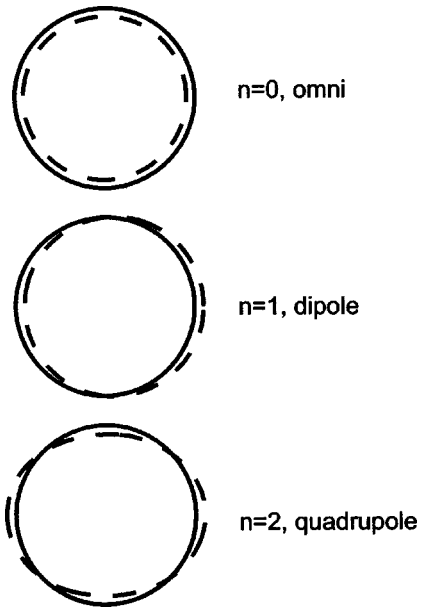


FIG. 1. Static (—) and deformed (---) displacements of omni, dipole, and quad modes.

motion of one half of the ring out of phase with the other and has a resonance frequency $f_1=\sqrt{2}f_0$. The third mode, with $n=2$, has four nodes and a resonance frequency $f_2=\sqrt{5}f_0$. The nth extensional mode has $2n$ vibrational nodes and a resonance frequency, $f_n=f_0(1+n^2)^{1/2}$. Figure 1 illustrates the first three even extensional modes of a ring.

This multimode transducer has application for directional acoustic links in an underwater acoustic modem system which requires compact directional acoustic transducers both for transmit and receive.[12] Underwater communication networks support telemetry of sensor data from distributed stations within the ocean, and are the basis for future undersea communication and navigation grids for autonomous undersea vehicles and submarines. The use of directional transducers substantially benefits network performance by increasing signal-to-noise ratio, decreasing multiaccess interference, reducing battery-energy consumption, and improving transmission security.[13]

The demand for compact directional transducers has revived interest in the use of multimode transducers. This paper describes a method for exciting a cylindrical transducer for operation as a directional multimode compact acoustic source and receiver which uses the first three extensional modes of vibration and operates in a band of frequencies between the dipole and quadrupole resonance frequencies.

II. BEAM PATTERN SYNTHESIS

Symmetrical beams are of particular interest and the analysis will be restricted to the even $\cos n\varphi$ components which requires an even excitation of the cylinder. The $\cos n\varphi$ radial motion of the nth extensional mode will generate a radiated acoustic pressure with a $\cos n\varphi$ azimuthal directional factor in a plane perpendicular to the axis of the cylinder. The fundamental, $n=0$, mode will generate an omnidirectional beam pattern, the $n=1$ mode will generate the two-lobed $\cos\varphi$ dipole beam pattern, and the $n=2$ mode will generate the four-lobed $\cos 2\varphi$ quadrupole beam pattern.

A theoretical model for the radiation from the cylinder may be based on the Laird–Cohen model[14] with infinite rigid extensions. In this case the azimuthal radial velocity distribution, u_r, on the cylinder surface may be expanded as the Fourier series

$$u_r=e^{-i\omega t}\sum_{n=0}^{\infty} u_n\cos(n\varphi). \tag{1}$$

We may also write the pressure, p, at a far field radial distance r as

$$p(r,\theta,\phi)=e^{-i\omega t}\sum_{n=0}^{\infty} p_n(r,\theta)\cos(n\varphi), \tag{2a}$$

where the modal value of the far field pressure is

$$p_n(r,\theta)=u_n 2\rho cL(e^{ikr}/\pi r)\mathrm{sinc}(kL\cos\theta)e^{-in\pi/2}/ H_n'(ka\sin\theta)\sin\theta. \tag{3a}$$

θ is the polar angle measured from the axis of the cylinder, φ is the azimuth angle, ρ is the density and c is the sound speed in the medium, k is the wave number, L is the half-length of the cylinder, a is the radius of the cylinder and $H_n'(ka\sin\theta)$ is the derivative of the cylindrical Hankel function of the first kind of order n. The length, $2L$, of the cylinder or array of cylinders determines the beam pattern in the vertical plane through the axis of the cylinder.

In the plane perpendicular to the axis of the cylinder $\theta=\pi/2$ and Eqs. (2a) and (3a) become

$$p(r,\phi)=e^{-i\omega t}\sum_{n=0}^{\infty} p_n(r)\cos(n\varphi), \tag{2b}$$

$$p_n(r)=u_n 2\rho cL(e^{ikr}/\pi r)e^{-in\pi/2}/H_n'(ka). \tag{3b}$$

Equation (3b) then allows us to write the nth order modal value in terms of the zero order omni mode value as

$$u_n/u_0=(p_n/p_0)e^{-in\pi/2}H_n'(ka)/H_0'(ka). \tag{4}$$

Since Eq. (2b) is a Fourier series we can obtain the modal values, p_n, from the far-field pressure, $p(r,\varphi)$, as

$$p_0(r)=(1/\pi)\int_0^{\pi} p(r,\phi)d\varphi, \tag{5a}$$

$$p_n(r)=(2/\pi)\int_0^{\pi} p(r,\phi)\cos(n\varphi)d\varphi \tag{5b}$$

for $n=0$ and $n>0$, respectively.

Thus, in theory, any specific directivity function or beam pattern can be synthesized through Eqs. (5a) and (b) with the required modal velocity ratios given through Eq. (4). A large number of p_n coefficients may be necessary if the desired beam pattern function $p(\varphi)$ changes abruptly or is narrow. In our case of interest we use only a three term expansion of $\cos n\varphi$ corresponding to the first three extensional modes. Here the pressure function expansion may be written as

$$p(\varphi)=p(0)(1+A\cos\varphi+B\cos 2\varphi)/(1+A+B), \tag{6}$$

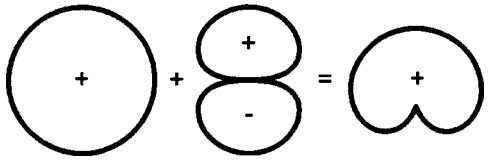


FIG. 2. The synthesis of a cardioid beam pattern from omni and dipole beam patterns.

where the normalized beam pattern function is $p(\varphi)/p(0)$. Substitution of Eq. (6) into Eqs. (5a) and (5b) yields the solutions $p_0/p(0)=1/(1+A+B)$, $p_1/p(0)=A/(1+A+B)$ and $p_2/p(0)=B/(1+A+B)$, which is also evident if the pressure function expansion is simply compared with Eq. (2) with $\theta=\pi/2$. Results can then be substituted into Eq. (4) to determine the required velocity ratios, u_n/u_0, for selected values of the weighting coefficients A and B.

Possibly the simplest directional beam pattern synthesis is the classic cardioid beam pattern. The cardioid directional factor is the superposition of an omni and a dipole directional factor, $\cos\varphi$, of equal pressure amplitude and phase so that $A=1$ and $B=0$. It has a strong null at its rear, as illustrated in Fig. 2, with corresponding normalized beam pattern function

$$p(\varphi)/p(0)=(1+\cos\varphi)/2. \tag{7}$$

The cardioid beam pattern is widely used in microphones and has been used in underwater acoustics[15] where unidirectional characteristics are desirable.

Adding higher order modes to the synthesis can widen the range of possible beam patterns. Adding omni, dipole, and quadrupole modes in the ratio 1:2:1 ($A=2$, $B=1$) gives a super cardioid beam pattern shown in Fig. 3. The corresponding beam pattern function is

$$\begin{aligned} p(\varphi)/p(0)&=(1+2\cos\varphi+\cos 2\varphi)/4 \\ &=\cos\varphi(1+\cos\varphi)/2 \end{aligned} \tag{8}$$

with the second equality revealing an alternative interpretation of the pattern as a product of dipole and cardioid beam patterns with nulls at $\pm 90°$ and 180°.

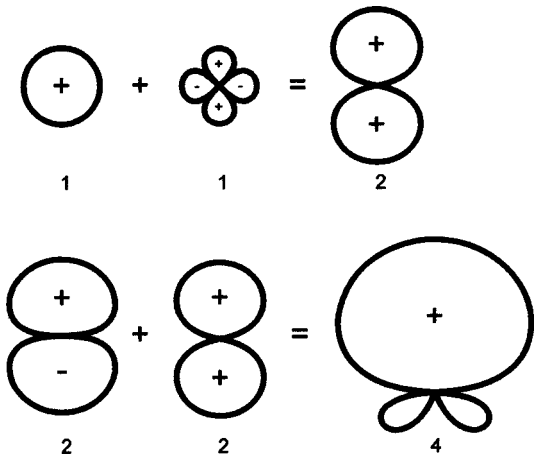


FIG. 3. The synthesis of a super cardioid beam pattern from omni, dipole, and quad beam patterns.

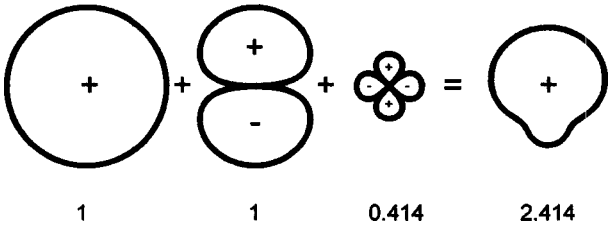


FIG. 4. The synthesis of a "quadrant" beam pattern.

A superposition with comparatively smaller dipole and quadrupole components, in the ratio 1:1:0.414, ($A=1$, $B=0.414$) is illustrated in Fig. 4 with the beam pattern function

$$p(\varphi)/p(0)=(1+\cos\varphi+0.414\cos 2\varphi)/2.414. \tag{9}$$

This combination of modes was chosen to give a pattern with a 90° beam width and a 15 dB front-to-back ratio as illustrated in Fig. 5. This pattern will be called the quadrant beam pattern; it is suitable for 45° incremental steering and is used in our implementation of the tri-modal directional transducer concept.

A difficulty comes in driving the cylinder to produce a superposition of pure extensional modes to generate a specific radial motion of the cylinder wall, with velocity coefficients u_n, and hence a desired beam pattern. Each extensional mode has a different resonance frequency and so, at any given frequency, some modes will be preferentially driven. The transfer function between modal electric field strength and modal velocity varies with frequency. In addition, at any given frequency the amplitude of the acoustic pressure generated by a modal velocity will depend on the mode order, with lower order modes radiating more effectively than higher order modes. Useful directional beam pattern synthesis is feasible, by using the low order modes in the vicinity of their resonances.

An appropriate distribution of voltages around the piezoelectric cylinder is required to excite the desired vibrational modes. A circumferential electric field voltage distribution $V(\varphi)=V_0\cos n\varphi$ will excite the $\cos n\varphi$ extensional mode. A continuous voltage distribution is difficult to implement but

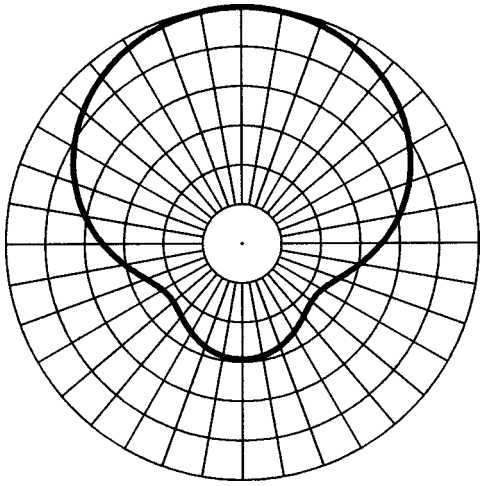


FIG. 5. Quadrant beam pattern distribution. Radial divisions are 5 dB, rotational divisions are 10°.

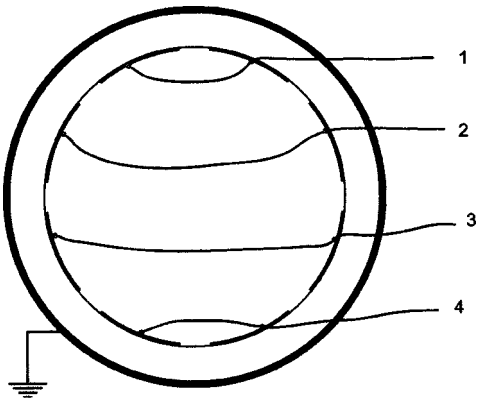


FIG. 6. A 31 mode piezoelectric ceramic cylinder with eight inner electrodes connected in pairs.

can be approximated by a set of equally spaced segmented inner electrodes with one electrode on the outside. An array of $2M$ electrodes can uniquely excite all even modes up to $n=M$. (The omni mode is a special case requiring only a single inner electrode.)

A particularly useful arrangement for the first three modes has eight electrodes. With even symmetry the eight electrodes can be grouped into four pairs as in Fig. 6. If the voltage distribution for the four pairs of electrodes is a four element vector, then the distribution $[1,1,1,1]$ will excite only the $n=0$ omni mode. As tabulated in Table I, the distribution $[1,1,-1,-1]$ will excite primarily the $n=1$ dipole mode while the distribution $[1,-1,-1,1]$ will excite primarily the $n=2$ quadrupole mode.

III. TRANSDUCER ANALYTICAL MODEL

If the motion of the cylinder is assumed to be the superposition of the set of extensional modes, an analytical model can be developed. The first part of the analysis is the calculation of the vibrational response to an applied voltage distribution. Although a finite element model was used in modeling the transducer, an approximate analytical model was also developed for comparison with the finite element model and for more rapid computation of transducer design variations. This model is based on the Fourier representation by Gordon, Parad, and Butler[16] for dipole mode operation of a piezoelectric ring, extrapolated to cover the case of higher order modes. Accordingly, for the *n*th mode and an applied even electrical field E_n of angular frequency ω, the radial velocity may be written as

$$u_n = j\omega\omega_0^2 a d_{31} E_n / [(1+n^2)\omega_0^2 - \omega^2 + (1-n^2\omega_0^2/\omega^2) j\omega Z_n/\rho_0 t], \tag{10}$$

where ω_0 is the fundamental angular resonance frequency, a is the mean radius of the cylinder, ρ_0 is the density of the ring, t is the wall thickness of the cylinder, and d_{31} is the piezoelectric coefficient. Z_n is the specific acoustic modal impedance, p_n/u_n, where p_n is the modal pressure. A more recent alternative model based on an energy representation has also been given by Aronov.[17]

TABLE I. Basic voltage distributions that excite each fundamental mode.

Mode	V_1	V_2	V_3	V_4	Sum
Omni	1	1	1	1	4
Dipole	1	1	−1	−1	0
Quadrupole	1	−1	−1	1	0

There is no analytical expression for either the modal specific acoustic impedance, Z_n, or the far field sound pressure amplitude for a finite unbaffled cylindrical transducer. However, there are useful approximations that can be made for an analytical model. The most accurate assumes that the transducer is embedded in an infinitely long rigid cylinder as modeled by Laird and Cohen.[14] In this case the Fourier series method by Butler and Butler[18] may be used to obtain the radiation impedance for each cylindrical mode. If the cylinder is long compared to the acoustic wavelength, the simpler model for an infinitely long cylinder[19] could also be used to obtain an approximation for Z_n. On the other hand, if the cylinder is acoustically small, a spherical model of the same surface area could be used. Although the Laird–Cohen model is probably the most accurate, we chose the spherical analytical model since our initial design was based on an acoustically small cylinder; and moreover, with this model algebraic expressions for the impedance as well as the far field pressure expansion could be used resulting in a rapid design program. The simple spherical model allowed many design iterations before selected specific designs were evaluated using a more accurate finite element model (FEM). As the design evolved, the transducer size became acoustically larger and the accuracy of the simple model became questionable. These analytical results were found, however, to be still of value and reasonably good when compared with the more accurate finite element results and the measured results within the operating band of interest.

An approximate model for the modal specific acoustic impedance for Eq. (10) may be obtained from a spherical wave expansion. If the surface of a sphere of radius a_s, polar angle θ, is vibrating with an axially symmetric radial velocity $u_r(\theta)$ then the complex amplitude of the acoustic pressure on the surface of a sphere can be expressed as the sum[19]

$$p(a_s,\theta) = \sum_{n=0}^{\infty} p_n P_n(\cos\theta), \tag{11}$$

where $p_n = b_n h_n(ka_s)$, P_n is the Legendre function of order n, and h_n is the spherical Hankel function of the second kind of order n. The coefficient b_n may be determined from the boundary conditions and the radial velocity on the surface of the sphere through

$$u_r(a_s,\theta) = \sum_{n=0}^{\infty} u_n P_n(\cos\theta), \tag{12}$$

where $u_n = -jb_n h_n'(ka_s)/\rho c$ was used to obtain Eq. (12).

The specific acoustic modal impedance for a sphere of radius a_s is then

$$Z_n = p_n/u_n = -j\rho c h_n(ka_s)/h_n'(ka_s). \tag{13}$$

With $x=ka_s$ we then get

$$Z_0 = \rho c[x^2+jx]/(x^2+1), \tag{14a}$$

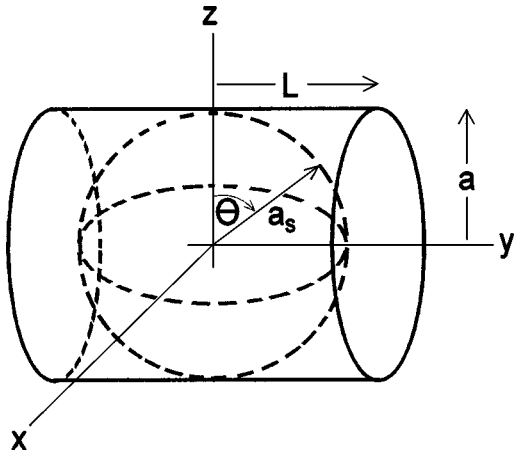


FIG. 7. Equivalent sphere of radius a_s within cylinder of radius a and half length L.

$$Z_1 = (\rho c/3)[x^4 + j(x^3 + 2x)]/(x^4 + 4), \tag{14b}$$

$$Z_2 = (\rho c/5)[x^6 + j(x^5 + 6x^3 + 27x)]/(x^6 - 2x^4 + 9x^2 + 81) \tag{14c}$$

for the specific acoustic modal impedance for the omni, dipole, and quadrupole modes of a sphere of radius a_s. Notice that the values of Z_n decrease as n increases. The approximate impedances for the multimode cylinder are obtained by equating the curved radiating areas of the cylinder, $4L\pi a$, and the sphere, $4\pi a_s^2$, yielding the equivalent sphere radius $a_s = (aL)^{1/2}$ where, as before, a is the radius and L is the half length of the cylinder. Equations (14a)–(c) when substituted into Eq. (10) yield the modal velocities, u_n, in terms of the electric field intensities, E_n.

Since the Legendre functions are polynomials in powers of $\cos\theta$, the spherical wave function expansion for the far field pressure as a function of the modal velocities can be expressed in terms of $\cos(n\theta)$. This expansion can then be written in an algebraic form and related to the modal expansion in $\cos(n\varphi)$ by aligning the symmetry axis, z, of the sphere with the y, z symmetry plane of the tri-modal cylinder, as illustrated in Fig. 7. This was done in our simplified analytical model which yields reasonably good agreement when compared to both a finite element model and measurements, satisfying the need for a rapid exploratory-design model.

IV. IMPLEMENTATION

In an earlier implementation of a tri-modal transducer, three different size cylinders were used in the form of a coaxial array.[20] The diameter of each cylinder was chosen to make the resonance frequency of the quadrupole mode, of the largest cylinder, and the dipole mode, of the next largest cylinder, coincide with the omni mode resonance frequency of the smallest cylinder. With this arrangement, each cylinder could be operated at resonance with approximately the same transmit voltage response. The voltage distributions required for a given beam pattern were then relatively easy to implement, with the required omni voltage distribution applied to the smallest cylinder, the dipole voltage distribution applied to the next largest, and the quadrupole voltage distribution applied to the largest. However, as the acoustic center of each cylinder was offset by its location in the array, these offsets and phase reversals on the back side of the cylinders created undesirable rear grating lobes in the superimposed beam patterns.

FIG. 8. A photograph of the encapsulated transducer with three piezoelectric cylinders and nine wire connector.

In the present implementation, shown in Fig. 8, the transducer is a stack of three cylinders of equal size. In this case each cylinder is driven with the same voltage distribution and the transducers are used in a frequency range between the resonance frequencies of the dipole and quadrupole modes, where the transmit voltage response of the first three modes are of the same order of magnitude.

The three cylinders are poled through the thickness of the wall. Each cylinder is 108 mm (4.25 in.) in diameter, 50.8 mm (2.00 in.) high with a wall thickness of 7.48 mm (0.19 in.). The outside of each cylinder has one continuous electrode, the inner surface has eight equally spaced electrodes around the circumference. The three cylinders are axially decoupled and wired in parallel to make an effective cylinder 152.4 mm (6.00 in.) high, encapsulated in polyurethane. The array height was chosen to yield a vertical beam width of approximately 28° for communication with other units approximately within the same horizontal plane. The diameter was chosen so the unit could be deployed from standard launch tubes. During evaluation, the eight inner electrodes were connected in pairs as shown in Fig. 6. This wiring scheme creates a plane of symmetry which contains the maximum response axis for the acoustic radiation. In actual practice each electrode is independently wired to allow steering.

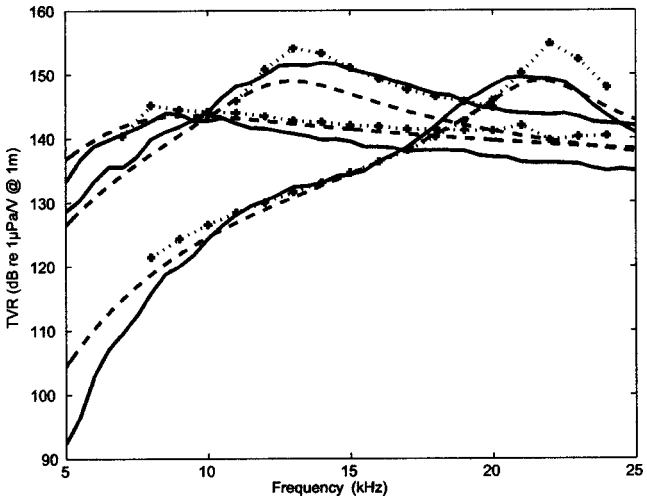


FIG. 9. Transmit voltage response for the transducer driven by the basic voltage distributions of Table I showing the measured (—), analytical (---), and FEM (+++) results.

Three basic voltage distributions, shown in Table I, were used to excite the transducer in each of the three fundamental extensional modes. Due to the discrete nature of the electrodes, the basic voltage distributions do not excite a single pure mode of vibration. The gaps between electrodes cause each distribution to excite small components of other modes as well as the targeted mode. For example the $n=1$ dipole mode voltage distribution also excites the $n=3$ mode to a small extent.

Figure 9 shows the measured transmit voltage response (TVR) for each of the voltage distributions of Table I. Also shown is the response predicted by the analytical model and by a finite element model. The in-water response for each voltage distribution shows that in each case, the desired mode has been excited. The omni mode resonates in the vicinity of 9 kHz, the dipole in the vicinity of 13 kHz, and the quadrupole in the vicinity of 22 kHz. The theoretical in-air resonances are 10.1, 14.3, and 22.6 kHz, respectively, indicating that the quadrupole mode experiences the least water mass loading, as might be expected from Eq. (14c). As previously mentioned, good agreement with the analytical model results should be limited to a frequency range where the transducer is small compared to the wavelength in the medium. The surprisingly reasonably good analytical TVR results, in the current higher range, are due, in part, to the weak dependence on the higher order radiation impedance values, especially when damping is included. The analytical model includes a component of mechanical loss which was adjusted to obtain 70% mechanical efficiency and match the measured omni mode results. No mechanical losses were included in the finite element model and so its predicted resonance peaks are higher than the values of the measured or analytical model.

In order to implement a given directional response, a voltage distribution is required that results in a superposition of the far field pressure response of each mode in the required relative amplitudes with the same phase. If V_o is the complex amplitude of the omni modal voltage and T_o is the complex amplitude of the transmit response, then the complex far field pressure, p_o, due to V_o is $p_o=T_oV_o$. Similarly, $p_d=T_dV_d$ is the pressure due to the dipole modal voltage, V_d, and $p_q=T_qV_q$ is the pressure due to the quadrupole voltage distribution. The transmit response is the product of two physical components, the transfer function from voltage distribution to cylinder wall velocity and the transfer function from wall velocity to the radiated sound field.

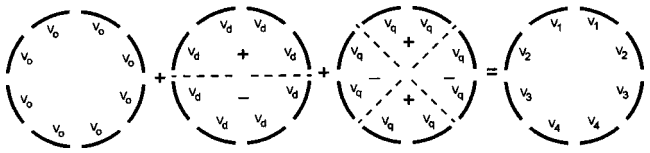


FIG. 10. An illustration of the addition of three basic monopole, dipole, and quadrupole voltage distributions to give the broadband voltage distribution.

With the desired beam pattern limited to the first three modes; the omni, dipole, and quadrupole modal voltages for the normalized directivity pressure function of Eq. (6) may be written as

$$V_o=1/T_o, \quad V_d=A/T_d, \quad V_q=B/T_q. \tag{15}$$

The coefficients $A=1$ and $B=0.414$ are for the specific case of the ''quadrant'' beam pattern, of Eq. (9), shown in Fig. 5.

The voltage distribution for the electrodes of Fig. 6 may be determined with the help of Fig. 10, which shows the phase reversals necessary for excitation of the dipole and quadrupole modes. This mode summation leads to the complex voltage distribution

$$V_1=V_o+V_d+V_q, \tag{16a}$$

$$V_2=V_o+V_d-V_q, \tag{16b}$$

$$V_3=V_o-V_d-V_q, \tag{16c}$$

$$V_4=V_o-V_d+V_q, \tag{16d}$$

for the four paired electrodes voltages, V_1, V_2, V_3, and V_4.

The voltage distribution over the cylinder for a required beam pattern can be found for any given frequency through Eqs. (16a)–(d). Since T_o, T_d, and T_q vary with frequency, V_o, V_d, and V_q and consequentially V_1, V_2, V_3, and V_4 must change with frequency to maintain the required beam pattern over a band of frequencies. Table II shows the values calculated from Eq. (15) for the omni, dipole, and quadrupole modes for the quadrant beam patterns at 15, 17.5, and 20 kHz. Table III shows V_1, V_2, V_3, and V_4, calculated from Eqs. (16a)–(d), at the particular frequencies.

It is possible, however, to find a fixed voltage distribution which approximately generates the desired quadrant beam pattern over the band. In the band of frequencies from 15 and 20 kHz, between the dipole and quadrupole resonances of Fig. 9, the transmit response of the first three

TABLE II. Calculated omni, dipole, and quadrupole modal voltages from Eq. (15).

Mode voltage	15 kHz	17.5 kHz	20 kHz
Omni V_o	$1.0+j0.0$	$1.0+j0.00$	$1.0+j0.00$
Dipole V_d	$0.1-j0.2$	$0.4-j0.10$	$0.55-j0.05$
Quadrupole V_q	$-1.3+j0.3$	$-0.6+j0.10$	$-0.25+j0.05$

TABLE III. Calculated voltage distributions from Eqs. (16a) to (d) for a quadrant beam pattern at three frequencies together with approximate constant broadband distribution.

Voltage	15 kHz	17.5 kHz	20 kHz	Broadband
V_1	$-0.2+j0.1$	$0.8+j0.0$	$1.3+j0.0$	$1.5+j0.0$
V_2	$2.4-j0.5$	$2.0-j0.2$	$1.8-j0.1$	$1.9+j0.0$
V_3	$2.2-j0.1$	$1.2+j0.0$	$0.7+j0.0$	$0.5+j0.0$
V_4	$-0.4+j0.5$	$0.0+j0.2$	$0.2+j0.1$	$0.1+j0.0$
Sum	$4.0+j0.0$	$4.0+j0.0$	$4.0+j0.0$	$4.0+j0.0$

modes are of the same order of magnitude. At low frequencies, the radial motion of the cylinder wall is in phase for all vibration modes, but as shown through the analytical results of Fig. 11, the low frequency far field transmit pressure response of the quadrupole mode leads the dipole by 90° and the dipole leads the omni by 90°. However, between 15 and 20 kHz, the omni and dipole modes have gone through their resonances and are approximately in phase while both are out of phase with the quadrupole. This observation allows the possibility of direct addition of the modes with only a 180° phase reversal needed for the quadrupole mode in the band between the dipole and quadrupole resonant frequencies.

The approximate broadband distribution of Table III was obtained from inspection of the distribution at the three separate frequencies mindful that the sum of each is and must be identically $4+j0$ since the dipole and quadrupole mode voltage distributions sum to zero, as illustrated in Table I. It was found that the distribution at 20 kHz yielded nearly satisfactory patterns at 15 and 17.5 kHz and that the imaginary parts were negligible. The distribution at 20 kHz was then modified to obtain nearly equal 90° beam patterns at 15 and 20 kHz with approximately 10 dB down level at 90° and 180°. Note that the selected broadband voltage distribution is purely real and that the distribution sums to 4. The single broadband distribution is a significant improvement in the implementation of the tri-modal transducer allowing one simple real voltage distribution (1.5,1.9,0.5,0.1) for the desired quadrant beam pattern over a band of frequencies from 15 to 20 kHz.

The broadband voltage distribution of the last column of Table III was implemented and the corresponding modeled and measured beam patterns, at four frequencies across the band, are shown in Figs. 12 and 13. A comparison of Fig. 12 with the ideal pattern of Fig. 5 shows that the simple broadband voltage distribution yields a slightly broader beam and a front to back ratio of 9 to 13 dB compared to the ideal value of 15 dB. The simple distribution yields a measured beam pattern which is approximately that of the quadrant directional beam at all four frequencies, with an approximate 90° beam width and a front to back ratio greater than 10 dB for all except the 15 kHz beam pattern. In operation this beam is steered in steps of 45° by incrementing the Table III broadband electrode voltage distribution. This is accomplished by incrementing the voltage amplitude on the eight amplifiers connected to the eight inner electrodes of the piezoelectric cylinder. No phase shifting is required. Figure 14 gives the calculated and measured transmit response for the broadband voltage distribution showing a nearly flat response across a band from 15 to 20 kHz.

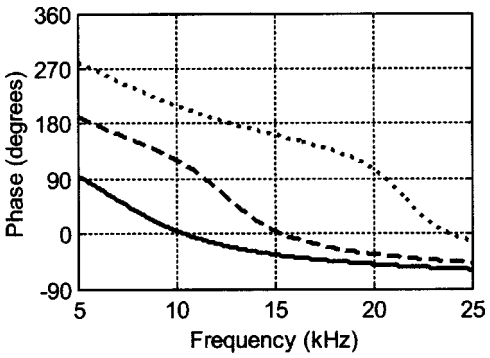


FIG. 11. Phase of the transmit response for the transducer driven by the basic voltage distributions of Table I for the omni (—), dipole (---), and quad (···) modes.

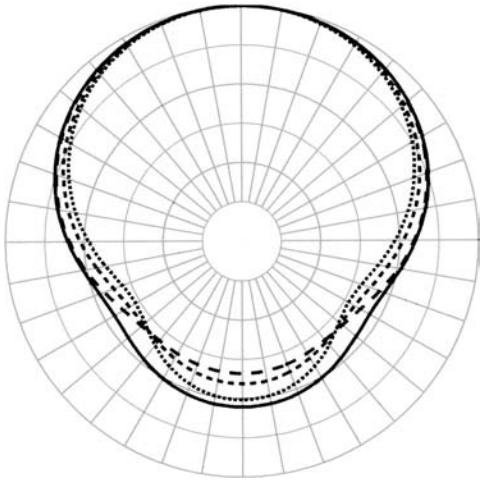


FIG. 12. Calculated beam patterns for the broadband voltage distribution of Table III at 15 kHz (—), 17 kHz (---), 19 kHz (---), and 20 kHz (···).

V. CONCLUSIONS

A simple, compact, easily steered, reasonably broadband directional sonar transducer has been implemented by ex-

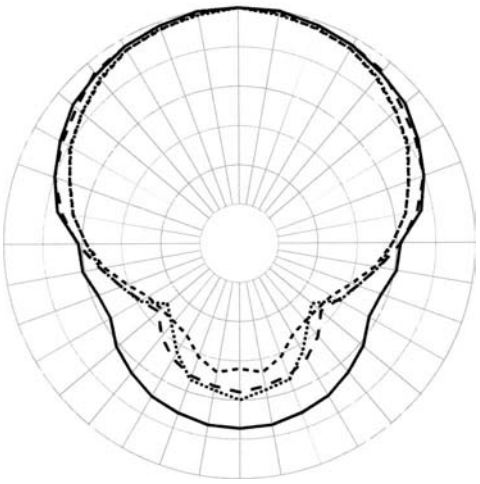


FIG. 13. Measured beam patterns for the broadband voltage distribution of Table III at 15 kHz (—), 17 kHz (---), 19 kHz (--) and 20 kHz (···).

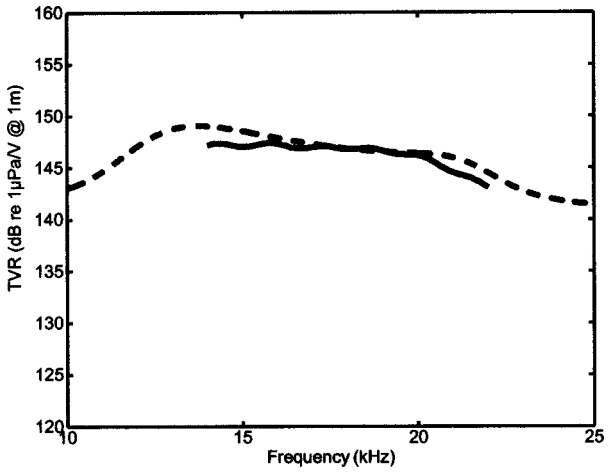


FIG. 14. Measured (—) transmit response for the broadband voltage distribution of Table III together with the prediction from the analytical model (---).

ploiting higher order extensional modes of a piezoelectric cylinder. The beam pattern responses of each higher order mode can be combined to produce a directional beam pattern. One particularly useful directional beam pattern is the quadrant directional beam formed by a superposition of the first three extensional modes of the cylinder (omni, dipole, and quadrupole modes with weightings 1:1:0.414). With segmented electrodes on a ceramic piezoelectric cylinder, a simple, easily implemented voltage distribution can be used to obtain nearly constant directional beam patterns over a significant band of frequencies. A transducer was fabricated, using this technique, which yielded a smooth transmit response from 15 to 20 kHz, with a nearly constant 90° beam width and a 10 dB front to back ratio across the band.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research through an SBIR Phase II contract. We would like to thank Bill Dalton and John Baker of FSI/Acoustikos, Inc., Cataumet, MA 02534, where the fabrication and measurements were performed. We would also like to thank Dr. Charles H. Sherman of Image Acoustics, Inc., for his helpful comments.

[1] B. Aronov, T. Oishi, L. Reinhart, and D. A. Brown, ''Broadband, multimode, free-flooded, baffled circular ring projectors,'' J. Acoust. Soc. Am. **109**, 2364 (2001).
[2] S. L. Ehrlich and P. D. Frelich, ''Sonar transducer,'' U.S. Patent 3,290,646, 6 December, 1966.
[3] J. L. Butler and S. L. Ehrlich, ''Superdirective spherical radiator,'' J. Acoust. Soc. Am. **61**, 1427–1431 (1977).
[4] S.-H. Ko, G. A. Brigham, and J. L. Butler, ''Multimode spherical hydrophone,'' J. Acoust. Soc. Am. **56**, 1809–1989 (1975).
[5] D. D. Ebenezer and P. Abraham, ''Piezoelectric thin shell theoretical model and eigenfunction analysis of radially polarized ceramic cylinders,'' J. Acoust. Soc. Am. **105**, 154–163 (1999).
[6] J. L. Butler, ''Directional flextensional transducer,'' U.S. Patent, 4,754,441, 28 June, 1988.
[7] S. C. Butler, J. L. Butler, A. L. Butler, and G. H. Cavanagh, ''A low-frequency directional flextensional transducer and line array,'' J. Acoust. Soc. Am. **102**, 308–314 (1997).
[8] J. Zhang, A. C. Hladky-Hennion, W. Jack Hughes, and R. E. Newham, ''A miniature class V flextensional cymbal; transducer with directional beam patterns: the double drive,'' Ultrasonics **39**, 91–95 (2001).
[9] H. F. Olson, *Acoustical Engineering* (Van Nostrand, New York, 1957), pp. 311–319.
[10] J. E. Cole III, *Acoustic Particle Velocity Sensors*, edited by M. J. Berliner and J. F. Lindberg [AIP Conf. Proc. **368**, 330–334 (1995)].
[11] A. E. H. Love, *Mathematical Theory of Elasticity*, 4th ed. (Cambridge University Press, London, 1934), p. 452.
[12] J. G. Proakis, E. M. Sozer, J. A. Rice, and M. Stojanovic, ''Shallow water acoustic networks,'' IEEE Commun. Mag. **39**, 114–119 (2001).
[13] J. T. Hansen, ''Link budget analysis for undersea acoustic signaling,'' Master's thesis, Naval Postgraduate School, Monterey, CA 93943, June, 2002.
[14] D. T. Laird and H. Cohen, ''Directionality patterns for acoustic radiation from a source on a rigid cylinder,'' J. Acoust. Soc. Am. **24**, 46–49 (1952).
[15] D. Stansfield, *Underwater Electroacoustic Transducers* (Bath University Press, Bath, 1990), p. 326.
[16] R. S. Gordon, L. Parad, and J. L. Butler, ''Equivalent circuit of a ring transducer operated in the dipole mode,'' J. Acoust. Soc. Am. **58**, 1311–1314 (1975).
[17] B. Aronov, ''Calculation of first-order cylindrical piezoelectric receivers,'' Sov. Phys. Acoust. **34**, 804–811 (1988).
[18] J. L. Butler and A. L. Butler, ''A Fourier series solution for the radiation impedance of a finite cylinder,'' J. Acoust. Soc. Am. **104**, 2773–2778 (1998).
[19] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw–Hill, New York, 1968).
[20] A. L. Butler, J. L. Butler, and J. A. Rice, ''Multimode directional telesonar transducer,'' J. Acoust. Soc. Am. **109**, 2363(A) (2001).

Effects of mutual impedance on the radiation characteristics of transducer arrays

Haksue Lee, Jinhee Tak, and Wonkyu Moon[a)]
Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), San 31, Hyojadong Namgu, Pohang, Kyungbuk 790-784, South Korea

Geunbae Lim
Bio Project Team, MEMS Laboratory, Samsung Advanced Institute of Technology (SAIT), San 14-1, Nongseo-ri, Kiheung-eup, Yongin-shi, Kyungki-do 449-712, South Korea

(Received 29 July 2002; revised 31 October 2003; accepted 6 November 2003)

The mutual resistance of transducer arrays is investigated in order to design arrays with improved performance for high intensity sounds at a given frequency. This work proposes the theory that the mutual resistance is related to the loading effects of pressure waves propagated from a piston driver on the surface of another driver. Using this interpretation, the important characteristics of the mutual resistance of two piston drivers are explained and the conditions for local maxima in the mutual resistance are easily determined. On the basis of analyses of the interactions between a driver and acoustic pressure waves, we propose a method to determine the driver radius and the distance between two drivers that give maximum mutual radiation resistance. To evaluate the proposed method, the total resistance of a transducer array is calculated using the formulas for mutual and self-resistance established by Pritchard. The results of the calculations of the total resistances of arrays with many drivers show that a transducer array with drivers arranged sparsely can achieve a larger value of the radiation power per unit area as well as better radiation efficiency than an array in which the drivers are in a closely packed arrangement at a given frequency. [DOI: 10.1121/1.1639328]

PACS numbers: 43.38.Hz, 43.20.Rz, 43.30.Yj [AJZ] Pages: 666–679

I. INTRODUCTION

Transducer arrays have been used in acoustic systems for a variety of purposes. For example, transducer arrays have been widely adopted in high power acoustic radiators to increase the output power and efficiency of acoustic radiation. One such application is the acoustic radiator for the low frequency active SONAR system, which uses a transducer array to increase radiation power and directivity in order to generate low frequency sounds through the nonlinear interactions of bi-frequency high-intensity acoustic beams.[1] Transducer arrays have also been employed in the high power generator in air such as the ultrasonic loudspeaker.[2,3] The aim when designing transducer arrays for such purposes is to assemble a collection of acoustic wave generation drivers such that the resulting array has enhanced output power due to the increase in the effective radiation area and, sometimes, due to the interactions between the drivers. The mutual impedance is a very important and useful concept in the design of arrays of drivers.[4,5] In the case of the direct radiation of low frequency sounds, it is almost impossible to achieve the required intensity of low frequency sounds with just one efficient acoustic driver; hence, a transducer array is constructed in such a way that the mutual radiation resistance between the drivers is as large as possible by arranging them as closely as possible.[4,6]

Transducer arrays are also utilized to produce high-intensity sounds in a relatively high frequency range in order to achieve two purposes: to increase output power by increasing the effective radiation area and to focus the emitted acoustic waves by properly arranging the drivers.[7,8] In general, the size of an acoustic driver is determined by a number of factors. For example, low-frequency-sound drivers should be large for efficient radiation, but the size of these drivers is limited to the maximum size that can be manufactured.[5] On the contrary, high-frequency-sound drivers are constrained to be small by structural characteristics such as inertial effects, natural frequencies, and the mode shapes of these frequencies.[8] As the size of a high-frequency driver is increased to greater than the optimal size, inertial effects increase to such an extent that its power efficiency decreases, and the radiation characteristics change according to the vibration modes of the structure, making it difficult to operate the driver. Therefore, it may be more convenient to use a transducer array to generate high-frequency sounds. In this case, the mutual impedance may provide important information on how to arrange the drivers in an array to obtain the maximum possible radiation efficiency and power. For relatively high operating frequencies, it is possible to build a unit driver such that the dimensionless parameter, ka $(=2\pi a/\lambda)$, is larger than 1.[7] Hence, the approach used to design transducer arrays for low-frequency radiators is inappropriate to be applied to the construction of transducer arrays for generating high-frequency sound. The optimal method to arrange the drivers in an array should be established by considering the radiation resistance of arrays for higher ka values. This approach is taken in the present study.

In this paper, we investigate the method for designing

a)Electronic mail: wkmoon@postech.ac.kr

transducer array patterns so as to achieve large radiation resistance by constructing an array of drivers of radius a, for radiating sounds of wave number k such that $ka>1$. Then, the power efficiency of a transducer array can increase if the unit drivers have high internal impedance. Even in the case that a transducer array is composed of unit drivers with sufficiently low internal dissipation and operating at its resonance frequency, a larger radiation resistance will, at least, provide a transducer designer with more freedom to improve the performance of an array while the power efficiency is maintained high. The approach taken here is to evaluate the radiation resistance of a transducer array over a wide frequency range by calculating the mutual and self-impedances based on the formula derived by Pritchard.[4] For convenience of calculation, the drivers in an array are assumed to be planar circular pistons on an infinite plane baffle.[4] Before calculating the impedances, the mutual effects are explained on the basis of interference effects. This approach helps to elucidate the trends in the mutual effects such as the oscillation of the mutual resistance with decreasing amplitude as the operating frequency increases observed in systems of two drivers with fixed driver size and center-to-center distance. The relationship between the frequency, the driver size, and the center-to-center distance between drivers is established for the mutual resistance of two drivers, after which the result is confirmed by calculating the mutual impedance using the formula of Pritchard. Further calculations showed that the trends in the mutual resistance of two drivers are similar to those in the mutual resistance for a number of drivers. From the analyses of the trends in the mutual resistance, it is discovered that proper selection of the separation of drivers in a transducer array can significantly improve the power efficiency or at least the output power of the array within a certain frequency range.

II. RADIATION IMPEDANCE CALCULATIONS USING MUTUAL IMPEDANCES

When assessing the performance of a transducer array, the mutual radiation impedance is a useful quantity because the radiation output power of the transducer array can be calculated using its radiation resistance at the operating frequency. Let R_r be the radiation resistance of an acoustic radiator. Then, its acoustic radiation power, W, is expressed as follows:[9]

$$W=R_r u^2, \tag{1}$$

where u is the particle velocity of the radiator surface. Thus, for a particular value of u, the output from the acoustic radiator is proportional to the magnitude of the radiation resistance. Hence, if the radiation efficiency of a radiator is defined as in Ref. 10, we may say that a larger value of radiation resistance gives more efficient radiation. But the power efficiency of a radiator is not guaranteed to increase as the radiation resistance of a radiator increases. The radiation efficiency is not identical to the power efficiency of a radiator because it does not contain the efficiency of electromechanical energy conversion inside a transducer.

The power efficiency of a radiator increases with the radiation resistance when the drivers of a transducer array have a much higher value of internal impedance than the acoustic radiation impedance. It is not in general true in the case of the radiators such as piezoelectric transducers with high Q factor in water operating at its resonance frequency. Even in the case of the radiators operated at the resonance frequency, the velocity control technique is widely used to guarantee a desired design goal on the performance of transducer arrays. If the velocity control technique is applied to a transducer array, it becomes a high impedance transducer because the velocity control technique makes internal impedance large by increasing electrical impedance of each driver in the array.[11] Therefore, it may be safely claimed that a more efficient transducer array can be designed by making the total radiation resistance larger. For these reasons, the radiation efficiency is considered instead of the power efficiency of a radiator in order to separate the effects of electromechanical transduction from those of acoustic radiation. And, throughout this paper, we say that sound radiation is efficient if radiation efficiency is high.

In this paper, for efficient sound radiation, the radiation impedance of a transducer array is investigated to choose the configuration of an acoustic driver array so as to give a large value of the mutual resistance. After understanding the characteristics of the radiation reactance and resistance of an array, the phase variations of acoustic waves emitted from the radiator can be predicted as the frequency varies.[9]

Consider a set of N identical circular planar piston drivers of radius a arranged on an infinite plane baffle. Assuming that all the pistons vibrate in the normal direction against the baffle with the same speed and phase, the total impedance of the piston array can be calculated as the follows:[4,12]

$$Z_t=\sum_{i=1}^{N}\sum_{j=1}^{N} Z_{ij}, \tag{2}$$

where

$$Z_{ij}=\rho c\pi a^2\sum_{p=0}^{\infty}\sigma_p(ka)\left(\frac{a}{d_{ij}}\right)^p\zeta_p^{(2)}(kd_{ij}) \quad \text{if } i\neq j, \tag{3}$$

$$Z_{ij}=\rho c\pi a^2[R_1(2ka)+jX_1(2ka)] \quad \text{if } i=j. \tag{4}$$

In these equations, d_{ij} is the distance between the centers of the ith and jth piston and k is the wave number. The functional formulas in Eqs. (3) and (4) are defined as follows:

$$R_1(x)=1-\frac{2J_1(x)}{x}=\frac{x^2}{2^2 1!2!}-\frac{x^4}{2^4 2!3!}+\cdots, \tag{5}$$

$$X_1(x)=\frac{4}{\pi}\left(\frac{x}{3}-\frac{x^3}{3^2 5}+\frac{x^5}{3^2 5^2 7}-\cdots\right), \tag{6}$$

$$\sigma_p(x)=\frac{2\Gamma(p+1/2)}{\pi^{1/2}}\sum_{n=0}^{p}\frac{J_{n+1}(x)J_{p-n+1}(x)}{n!(p-n)!}, \tag{7}$$

$$\zeta_p^{(2)}(x)=(\pi/2x)^{1/2}[J_{p+1/2}(x)+j(-1)^p J_{-p-1/2}(x)]. \tag{8}$$

Here, $J_n(x)$ is Bessel function of the first kind and Γ is gamma function. From the above-mentioned formulas, the total resistance, R_t, and total reactance, X_t, can be derived as follows:

$$R_t = \mathrm{Re}[Z_t] = \rho c \pi a^2 \left\{ \left[\sum_{i=1}^{N} \sum_{\substack{j=0 \\ i \neq j}}^{N} \sum_{p=1}^{\infty} \sigma_p(ka) \times \left(\frac{a}{d_{ij}}\right)^p \left(\frac{\pi}{2kd_{ij}}\right)^{1/2} J_{p+1/2}(kd_{ij}) \right] + N \cdot R_1(2ka) \right\}, \quad (9)$$

$$X_t = \mathrm{Im}[Z_t] = \rho c \pi a^2 \left\{ \left[\sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \sum_{p=0}^{\infty} \sigma_p(ka) \times \left(\frac{a}{d_{ij}}\right)^p \left(\frac{\pi}{2kd_{ij}}\right)^{1/2} (-1)^p J_{-p-1/2}(kd_{ij}) \right] + N \cdot X_1(2ka) \right\}. \quad (10)$$

In Eq. (9), the first term including the summation function represents the mutual resistance and the second term represents the self-resistance. By the same rule, the first term in Eq. (10) can be described as the mutual reactance and the second term can be called the self reactance.

The above-presented equations for calculating the impedance of a circular piston array [Eqs. (2)–(10)], which were formulated by Pritchard, are valid if there are no scattering effects or if the size of the radiator is much smaller than the acoustic wavelength.[4] If the scattering effect is not negligible, as is the case for the drivers of a spherical radiation surface, the following condition should be satisfied: $ka \leqslant 1$.[13] In the case of a plane circular piston driver, however, the scattering effects are negligibly small even at relatively high values of ka.[6,14] Therefore, Eqs. (2)–(10) can be used to calculate the total impedance of an array of piston drivers.

When identical planar circular pistons of radius a are arranged on a planar infinite baffle such that they are separated by a center-to-center distance of d_{ij}, the radiation impedance of the resulting array can be easily calculated using Eqs. (2)–(10). In the numerical calculation of the radiation impedance of a specific array, the infinite series in Eqs. (9) and (10) should be approximated by a sum over sufficient terms to obtain the desired level of accuracy. Since the pistons cannot be physically overlapped, the value of a/d_{ij} is always smaller than 1/2 and consequently the infinite series can be approximated to a numerical accuracy of less than 1% by taking only the first six terms. Furthermore, careful comparison of the results obtained using only the first four terms with those obtained from calculations including a greater number of terms showed that there is no significant difference between the results obtained using four terms and those obtained using a greater number of terms. Therefore, only the first four terms are used to calculate the self- and mutual resistance in the present study.

III. OPTIMAL DISTANCE BETWEEN A PAIR OF DRIVERS BASED ON MUTUAL EFFECTS

A. Mutual effects between two pistons

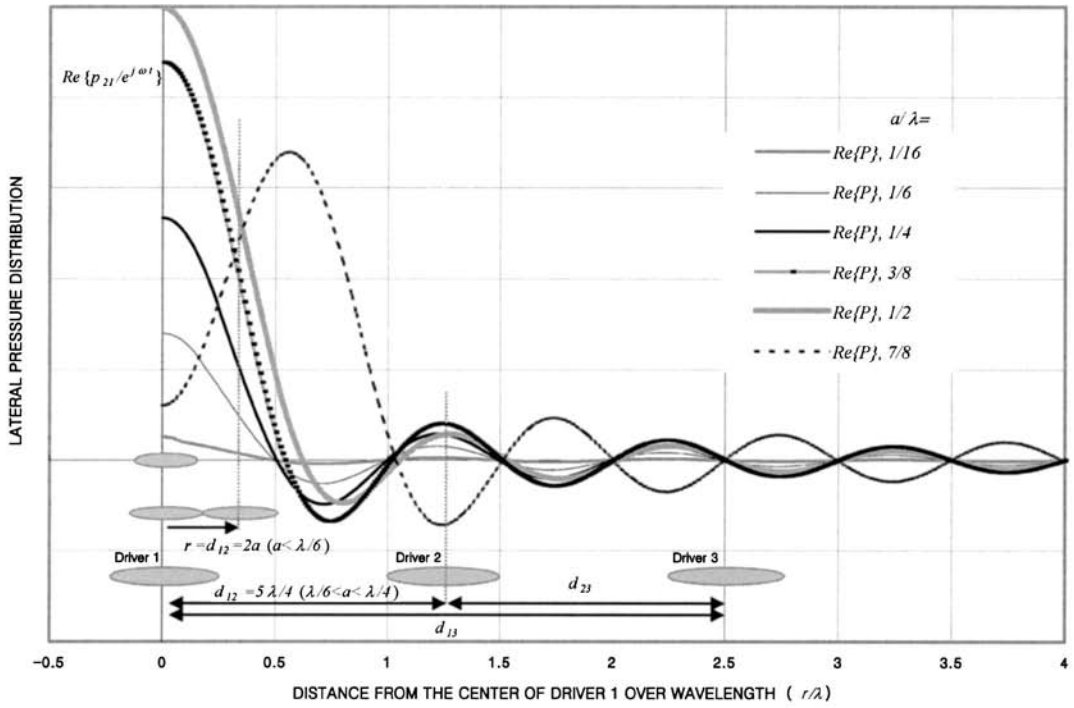
When arranging several pistons, we should first consider the mutual effects between two pistons. If the acoustic wave from one driver has a loading effect on the other, the radiation resistance of the latter increases. Therefore, if the distance between the neighboring drivers is determined so that they affect each other to increase the loading effects, the optimum distance between drivers can be defined as the distance at which the radiation resistance is maximized. For one driver (Driver 1) to have the maximum positive loading effect on a neighboring driver (Driver 2), the acoustic pressure wave generated by Driver 1 should be in phase with the velocity of Driver 2.[14–16] The reason for this is that the acoustic pressure propagated from Driver 1 exerts resistive forces on the surface of Driver 2 if it is in phase with the velocity of Driver 2 when it arrives on the surface of Driver 2. Then, the output power of Driver 2 should increase due to the loading by the pressure wave propagated from Driver 1. This argument can be applied to a system of two circular piston drivers moving in phase, which is considered in the remainder of this section.

Figures 1(a) and (b) show the spatial distribution of the real part and phase of the acoustic pressure divided by the time-dependent harmonic term ($p/e^{j\omega t}$) radiated from Driver 1 in the radial direction when it is vibrating harmonically ($u_1 = U_1 e^{j\omega t}$) in the direction perpendicular to the planar baffle where ω is the angular frequency of the velocity of piston driver.[12] Let the wavelength of the acoustic wave at the given frequency, ω, be λ. The acoustic pressure at time t and location r can be expressed as $p = P_0(r)\cos[\omega t + \phi(r)]$, and the velocity of Driver 2 is $u = U_0 \cos(\omega t)$, which is identical to that of Driver 1. Then, the loading effects on Driver 2 by the acoustic pressure wave propagated from Driver 1 can be expressed at location r as follows:[12]

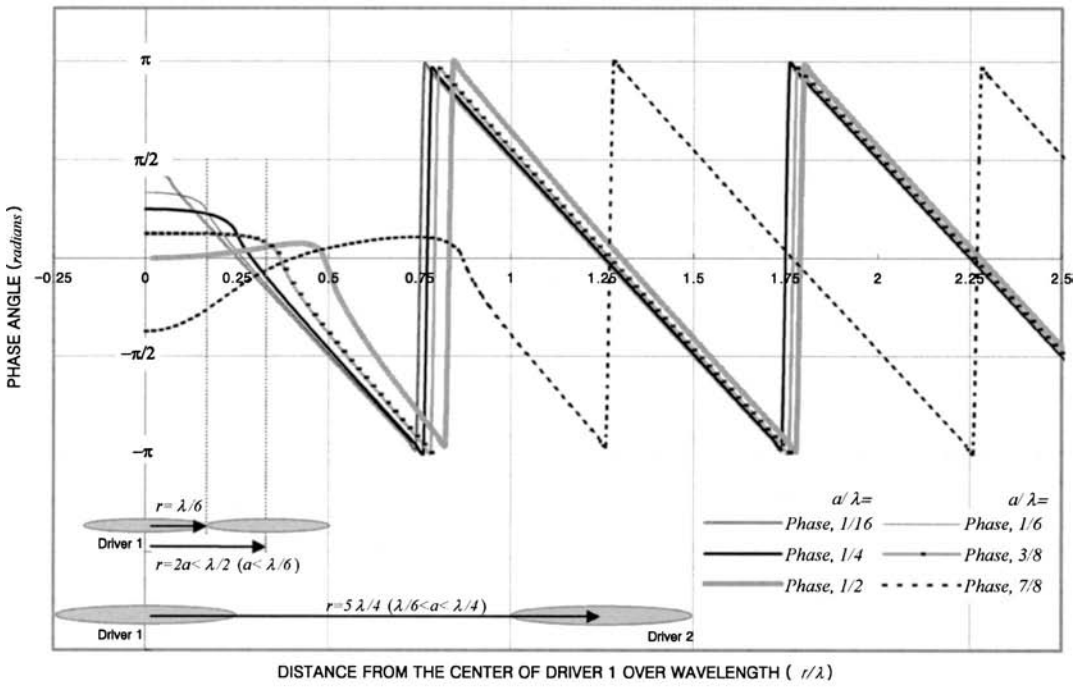
$$W = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} P_0(r)\cos[\omega t + \phi(r)] U_1 \cos(\omega t) dt = \frac{U_1 P_0(r)}{2} \cos[\phi(r)] = \frac{U_1}{2} \{P_0(r)\cos[\phi(r)]\}. \quad (11)$$

Hence, the loading effect is greatest when $\phi(r) = 0$ and zero when $\phi(r) = \pi/2$. Additionally, the loading effect should be positive when $-\pi/2 < \phi(r) < \pi/2$, and negative when $-\pi < \phi(r) < -\pi/2$ or $\pi/2 < \phi(r) < \pi$. Since the real part of $p/e^{j\omega t}$ is $P_0(r)\cos[\phi(r)]$, which determines the r-dependent characteristics of the loading W shown in Eq. (11), the real part of $p/e^{j\omega t}$ represents the loading effects on Driver 2 due to the acoustic pressure wave propagated from Driver 1. In the following discussion of the mutual effects of two drivers, Fig. 1(a) will be used to measure the loading effects.

The graphs in Fig. 1(a) show several notable characteristics. If the ratio of the piston radius to the wavelength, $a/\lambda (= ka/2\pi)$, is smaller than 1/4, the loading effect becomes zero at locations where the distance from the center of Driver 1, r, is $\lambda/2$, λ, $3\lambda/2$, 2λ, $5\lambda/2$, etc., and it reaches its positive or negative peaks at $r = 3\lambda/4$, $5\lambda/4$, $7\lambda/4$, $9\lambda/4$, etc. In the region where $r < 2\lambda$, this periodicity in the locations of zeros and peaks becomes obscure when a/λ becomes larger than 1/2, while the locations of zeros and peaks are almost periodic in the region $r > 2\lambda$. When $1/4 < a/\lambda < 1/2$,



(a)



(b)

FIG. 1. The spatial distribution of the acoustic pressure radiated from Driver 1 in the radial direction when it is vibrating harmonically: (a) real part of the pressure distribution, (b) phase variation.

the locations of the zeros and peaks are very close to the locations of those when $a/\lambda \leqslant 1/4$, even in the region $r < 2\lambda$.

Now consider the mutual effects between two drivers, Driver 1 and Driver 2, for systems in which $a/\lambda \leqslant 1/4$. If the distance between the centers of Drivers 1 and 2 is $\lambda/2$, the pressure wave generated from Driver 1 exerts a positive loading effect on about half of the surface of Driver 2 (the part of the surface separated from the center of Driver 1 by a distance of less than $\lambda/2$), while the pressure wave from Driver 1 exerts a negative loading effect on the remaining half of the surface of Driver 2 (the part that is separated from

the center of Driver 1 by a distance greater than $\lambda/2$). Thus, when the center-to-center distance is $\lambda/2$, the acoustic wave propagated from Driver 1 does not exert any significant loading effect on Driver 2 because the positive and negative loadings on each half of the radiation surface cancel each other such that the average loading is negligible. This effect leads to zero loadings when the center-to-center distance between Drivers 1 and 2 is $n\lambda/2$, where n is a positive integer. In contrast, when the driver centers are separated by $3\lambda/4$ the acoustic wave generated from Driver 1 exerts a negative loading effect over the entire surface of Driver 2. At this separation, the loading effect should reach its negative peak. This negative loading effect occurs when the center-to-center distance between Drivers 1 and 2 is $(4n-1)\lambda/4$, where n is a positive integer. The maximum positive loading effect is predicted to occur when the center-to-center distance between Drivers 1 and 2 is $(4n-3)\lambda/4$, where n is a positive integer, because at these values of the separation the entire surface of Driver 2 is inside the positive zone of the pressure wave propagated from Driver 1.

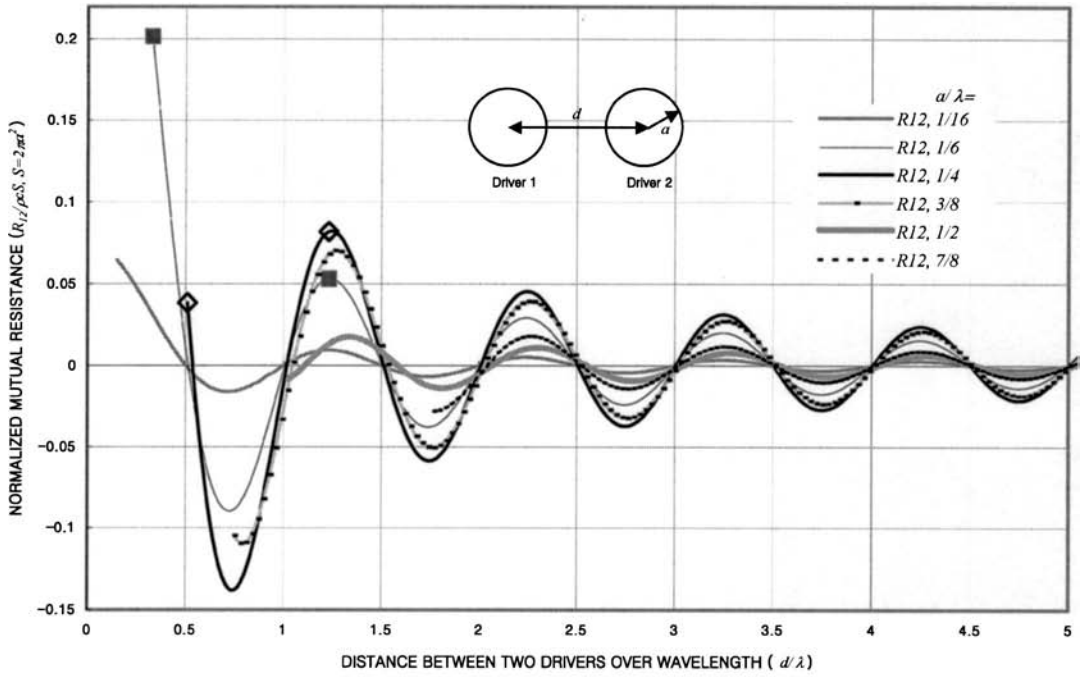
However, when a/λ is greater than 1/2, the above-presented arguments for the case of $a/\lambda \leqslant 1/4$ become difficult to apply to the situations illustrated in Fig. 1(a), primarily because the periodicity in the locations of the zeros, negative peaks, and positive peaks of the wave forms becomes obscure, and partly because the entire surface of Driver 2 cannot be included inside the zone of positive or negative acoustic pressure. However, even for the condition that $a/\lambda > 1/2$, the periodic tendency in the zero and peak loading effects can be predicted when the center-to-center distance is greater than 2λ, even though the exact center-to-center distances cannot be predicted using this type of analysis. When a/λ is between 1/4 and 1/2, the center-to-center distances for the zero and positive peak loading effects would be close to their values when $a/\lambda \leqslant 1/4$, even in the region from λ to 2λ. This is because over this range of r/λ the zero and peak locations of the acoustic pressure wave are almost periodic and almost all of the area of the Driver 2 surface can be included inside a second positive loading region of equal sign.

Now we extend the above-presented arguments to more specific conditions. When $a/\lambda \leqslant 1/6$, the acoustic radiation resistance is predicted to reach its maximum when Driver 2 is located as close as possible to Driver 1, i.e., when the center-to-center distance equals the sum of the radii of the drivers. This is because the entire surface of Driver 2 can be included inside the largest first positive peak region in Fig. 1(a). Therefore, as the center-to-center distance between Drivers 1 and 2, d_{12}, becomes smaller, the amplitude of the pressure wave loading on Driver 2 increases. This conclusion, that the theoretical optimum distance is $2a$, is the same as that reached by Pritchard from calculations of the mutual impedance of two drivers when ka is much smaller than 1.[4,16]

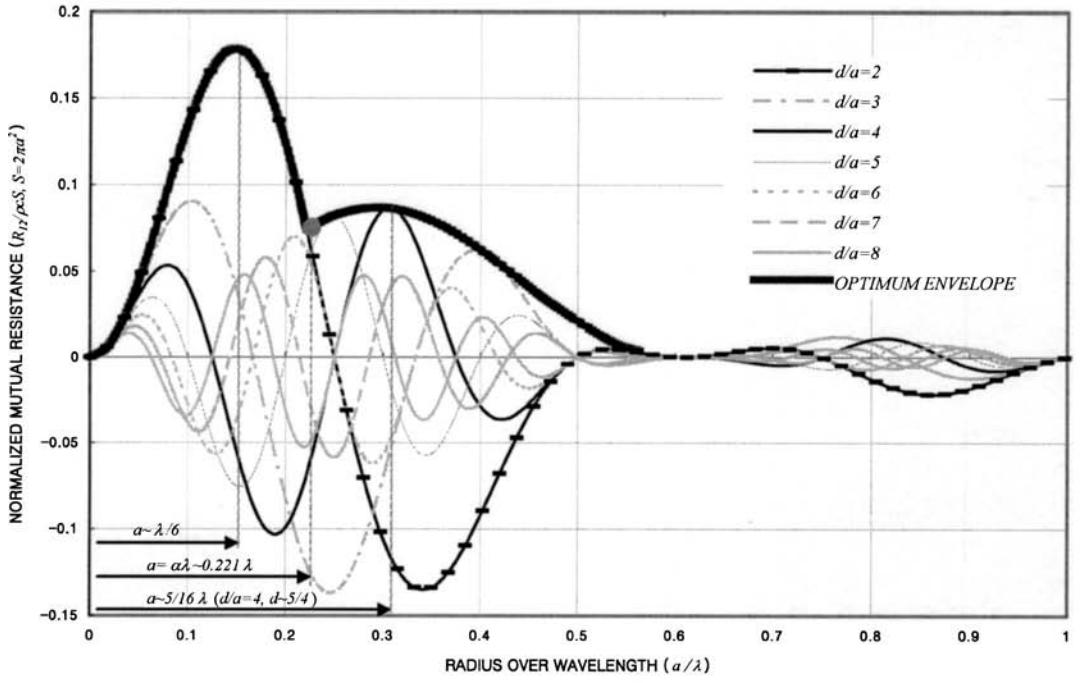
When a is greater than $\lambda/6$, however, Driver 2 becomes partially exposed to the region of negative loading effects even when it is as close as possible to Driver 1 (i.e., $d_{12} = 2a$). This exposure to the negative loading region reduces the resistive loading on Driver 2 caused by the acoustic wave from Driver 1. When $1/6 < a/\lambda \leqslant 1/4$, since the entire surface of Driver 2 can be included in the second positive loading zone around $d_{12} = 5\lambda/4$ of the pressure wave from Driver 1 [See Fig. 1(a)], increase in the loading effects on Driver 2 by the pressure wave does not diminish considerably if $d_{12} = 5\lambda/4$. If another driver, Driver 3, is located on the same line such that its center-to-center distance from Driver 2 is $d_{23} = 5\lambda/4$, the positive loading effect on Driver 2 increases due to Driver 3, but Driver 1 receives almost no loading effect from Driver 3 because the center-to-center distance between Drivers 1 and 3, d_{13}, is $10\lambda/4 = 5\lambda/2$, a distance at which almost no loading effect is predicted by the above-presented arguments. However, if the driving frequency is changed from c/λ to c/λ_1, where $\lambda_1 = 10\lambda/9$, Driver 3 located at $r = 10\lambda/4$ with respect to Driver 1 moves into the positive loading region of the acoustic pressure wave at frequency c/λ_1 from Driver 1. The important fact here is that the frequency c/λ_1 is close to c/λ (the difference is only $0.1c/\lambda$). Therefore, the wave propagated from Driver 3 will provide a positive loading effect of considerable magnitude on Driver 1 at the frequency c/λ_1, which is close to the frequency c/λ at which Driver 2 adds the maximum positive loading effects on Driver 1. This can explain the shifts of the peaks and zeros in the case of arrays of multiple drivers. In addition, even at the frequency c/λ, the positive loading effects should be slightly larger than the negative ones when $r = 10\lambda/4$ since the amplitude of the positive peak adjacent to the location $r = 10\lambda/4$ is larger than that of the adjacent negative peak, as can be seen in Fig. 1(a). Therefore, the mutual effects may increase with the number of drivers in the array.

Figure 2(a) shows the result obtained from calculating the mutual radiation resistance between two identical pistons using Eq. (9). In this figure, each curve illustrates the dependence of the mutual resistance on the distance d between Drivers 1 and 2, for a fixed piston radius, a, and frequency, k. The data show a trend similar to that found in the analysis presented earlier: The loading effects on the drivers vary with d in a somewhat periodic fashion. When a is less than $\lambda/4$, the peaks of the mutual resistance are observed at periodic values of d ($5\lambda/4$, $9\lambda/4$, etc.). When a is greater than $\lambda/4$, the periodicities in the mutual resistance are obscured; and when a is smaller than $\lambda/6$, the mutual resistance takes on a maximum value when the drivers are in the closest possible proximity ($d = 2a$). When a is equal to $\lambda/4$, the maximum positive mutual resistance is achieved at $d = 5\lambda/4$. Therefore, there is a value of a, say $\alpha\lambda$, larger than $\lambda/6$ and smaller than $\lambda/4$ (i.e., $1/6 < \alpha < 1/4$), where the maximum mutual resistance attained under the condition $d = 2a = 2\alpha\lambda$ becomes identical to that under the condition $d = 5\lambda/4$ [see Fig. 2(b)]. The numerical calculations give the result that $\alpha \approx 0.221$. Most previous studies have concluded that the maximum mutual effect is achieved when d/a is as close to 2 as possible.[4,16] The present results lead to a slightly different conclusion: if $a > \alpha\lambda$, the mutual effects are greater when d is around $5\lambda/4$ than when d is close to $2a$, because the conclusions of the previous studies can be applied only to cases in which ka is smaller than 1 (i.e., $a < \alpha\lambda$).

When investigating the influence of driver size on mutual impedance, it is helpful to plot graphs of mutual imped-



(a)



(b)

FIG. 2. Normalized mutual resistance between two piston drivers array: (a) mutual resistance for several frequencies (a/λ) varying the distance between two drivers (d/λ), (b) mutual resistance for several d/a values varying the frequency (a/λ).

ance against a/λ at constant d/a. A set of such graphs is shown in Fig. 2(b). It is evident from these plots that the maximum mutual effect is achieved at $d/a=2$ when $a<\alpha\lambda$, and at $d\approx 5\lambda/4$ when $\lambda/2>a>\alpha\lambda$. When the radius of the unit driver is greater than $\lambda/2$, the mutual effect is significantly reduced. Moreover, if the size of the unit driver can be arbitrarily reduced, the maximum value of the mutual effects is found when $a\approx\lambda/6$. Since the self-resistance of each driver increases significantly with increasing a/λ, the most compact arrangement of drivers does not guarantee the highest radiation efficiency in systems where the radius of the unit driver can be arbitrarily chosen.

B. Driver arrangement considering mutual effects

On the basis of the conclusions drawn in Sec. III A, we now consider the method for arranging an array of acoustic

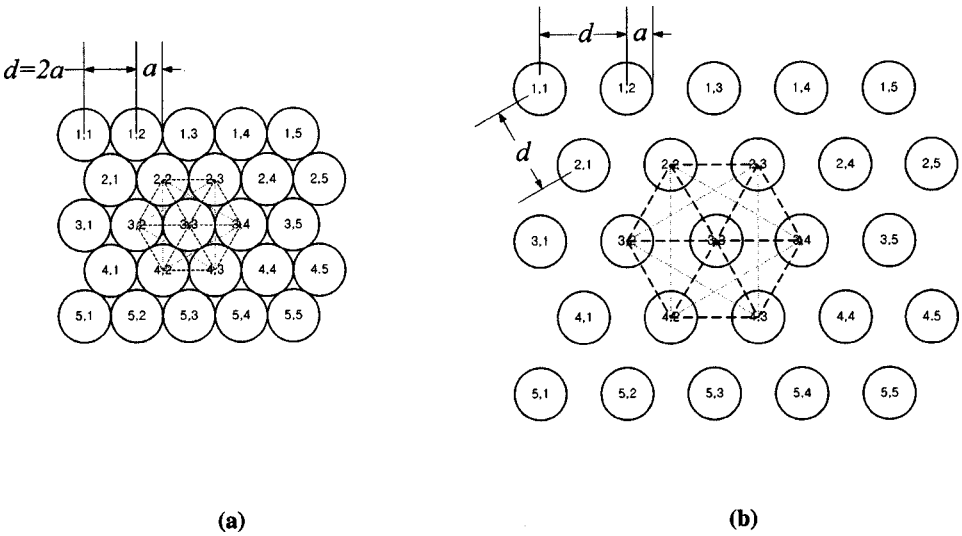


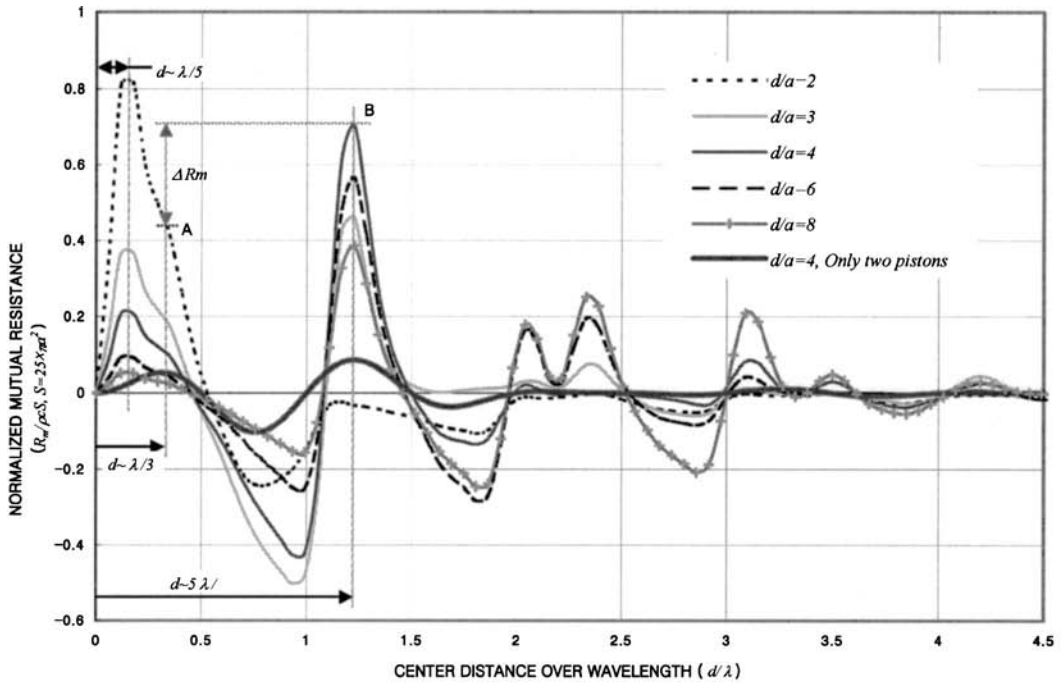
FIG. 3. Two hexagonal array patterns: (a) compact hexagonal array ($d/a=2$, $5\times5=25$ units in an array), (b) spaced hexagonal array ($d/a>2$, $5\times5=25$ units in an array).

circular piston drivers for use as a high-power acoustic radiator operating at relatively high frequency. First, we know from the above-presented analysis that mutual effects are significantly reduced if the size of the unit driver exceeds half of the wavelength of the given operating frequency (i.e., $a>\lambda/2$); hence, if mutual effects are to be used to achieve higher radiation efficiency, the radius of the unit driver, a, should be less than half of the wavelength of the given operating frequency ($a<\lambda/2$). The above-presented analysis also indicated that when the radius of the unit driver, a, is smaller than $\alpha\lambda$, the most compact arrangement of drivers provides the maximum mutual resistance. In this case, maximum mutual resistance is obtained when $a\approx\lambda/6$, as shown in Fig. 2(b). However, these optimal parameters apply only to the arrangement of two drivers. For arrangements of more than two drivers, the mutual effects due to additional drivers must be considered.

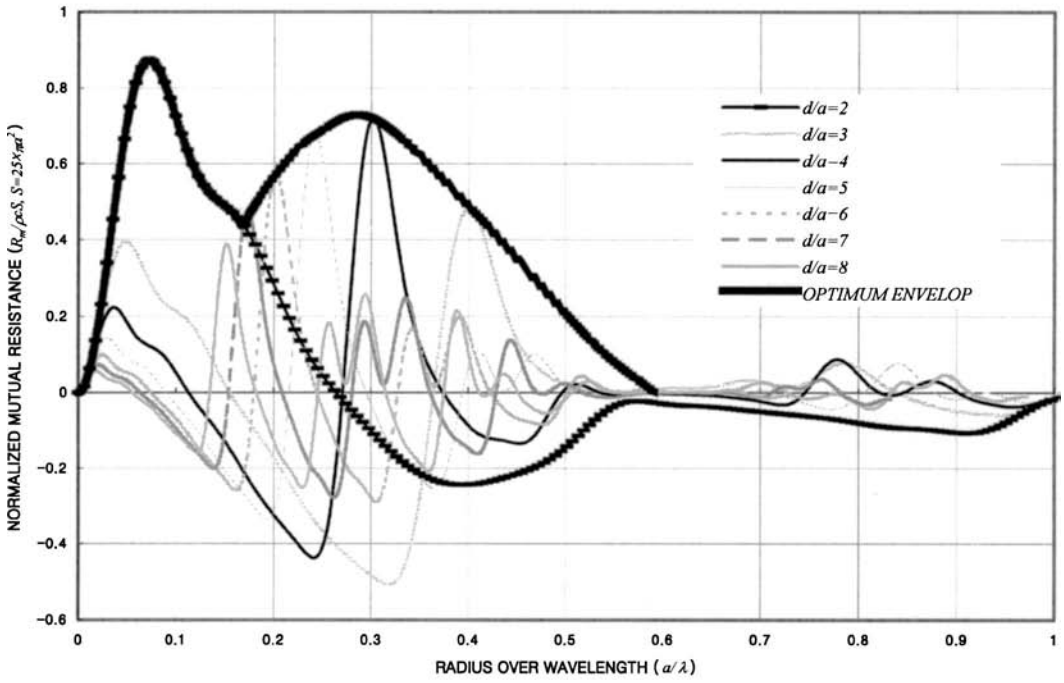
As shown in Sec. III A, in cases where the radius of the unit driver, a, can be made longer than $\alpha\lambda$, the mutual effects can be maximized by locating Driver 2 a distance of $5\lambda/4$ from Driver 1. For this center-to-center distance, the value of a that gives the maximum mutual resistance can be found by reference to Fig. 1(a). In Fig. 1(a), which shows data for only the selected six values of a/λ, the optimal value of a appears to be $3\lambda/8$, in which case the amplitude of in-phase acoustic pressure is the highest at $d=5\lambda/4$. Since the peaks when $a=\lambda/4$ and $a=3\lambda/8$ are almost the same, it is difficult to determine whether the maximum acoustic pressure amplitude at $d=5\lambda/4$ lies in the region $\lambda/4\leqslant a\leqslant3\lambda/8$ or in the region $3\lambda/8\leqslant a\leqslant\lambda/2$. However, careful inspection reveals that the peaks lie at values of d greater than $5\lambda/4$ when the radius is in the range $3\lambda/8\leqslant a\leqslant\lambda/2$. Therefore, the maximum loading effect is expected to occur in the region $\lambda/4\leqslant a\leqslant3\lambda/8$. In Sec. IV A it will be shown by calculating the mutual resistance of an array of multiple drivers that the maximum mutual resistance can be achieved with a larger than $\lambda/4$ and smaller than $3\lambda/8$, if d is set to be $5\lambda/4$.

Now consider the arranging of drivers into an array. As indicated by Pritchard,[4,16] the arrangement that maximizes the number of drivers adjacent to a driver while packing the drivers as compactly as possible (i.e., $d=2a$) is the hexagonal pattern, as illustrated in Fig. 3(a). In this pattern, the centers of adjacent drivers are located at the apexes of a hexagon centered on a given driver. Hexagonal arrays of drivers can also be constructed when the central distance is greater than $2a$, as shown in Fig. 3(b). Since the maximum mutual effects can be achieved when $d=5\lambda/4$ and $\lambda/4\leqslant a\leqslant3\lambda/8$, the optimal arrangement will be one in which as many adjacent drivers as possible are located around a given driver such that $d=5\lambda/4$ for every pair of adjacent drivers. Therefore, the hexagonal configuration with $d=5\lambda/4$ as the nominal center distance between drivers should be the ideal arrangement.

Having selected a hexagonal configuration as the basic arrangement for the transducer array, the question remains as to the best approach for maximizing radiation efficiency at a given operating frequency: a compactly arranged array of smaller drivers, or an array of larger drivers with a center-to-center distance of $d=5\lambda/4$. For a system of only two drivers, the optimal choice is a compact arrangement of drivers with radii of $a=\lambda/6$. However, in the case of a hexagonal array of many drivers, it is difficult to determine the optimal pattern of drivers (d/a) and the optimal value of a/λ because the use of a unit driver of radius $a=\lambda/6$ cannot guarantee maximum mutual effects. Consider a hexagonal array of drivers of radius $a=\lambda/6$ in which adjacent drivers are separated by a center-to-center distance of $d=2a$. As discussed in Sec. III A, the mutual resistance between two neighboring drivers has a large positive value. However, the mutual resistance between a given driver and a second neighbor driver ($d=4a$) is a negative number of fairly large magnitude [see Fig. 2(b)]. The mutual effect due to the third neighbor is almost zero, while that due to the fourth neighbor is positive but of much smaller amplitude than that due to the second neighbor. In addition, because the self-resistance of a unit driver decreases with decreasing size of the driver, the total resistance of an array of a number of drivers is very difficult to predict.



(a)



(b)

FIG. 4. Normalized mutual resistance of hexagonal array (5×5=25 units in an array): (a) mutual resistance for several d/a values along the d/λ axis, (b) mutual resistance for some d/a values along the a/λ axis.

IV. TOTAL RADIATION IMPEDANCE OF A HEXAGONAL ARRAY

A. Mutual resistance of a hexagonal array

Here, the mutual resistances of hexagonal arrays of 25 (5×5) transducer drivers are calculated using Eq. (9) and plotted in detail to investigate their radiation characteristics. Figure 4(a) shows plots of the mutual resistance versus d/λ for 5×5 arrays characterized by various values of d/a. As in the graphs for the two driver system [Fig. 2(a)], the graphs for the hexagonal arrays show a common peak in the mutual resistance at $d=5\lambda/4$. In the region of $d<\lambda/2$, all of the hexagonal arrays considered in Fig. 4(a) show additional peaks. Compared with the graphs for arrays of two drivers [Fig. 2(a)], the locations of the peaks in the region $d<\lambda/2$ are shifted toward the origin [Fig. 4(a)]. The peaks in this region are located at $d=\lambda/5$, which implies that $a\leqslant\lambda/10$,

whereas the corresponding peaks are located at $d=\lambda/3$ in the case of an array of two drivers. The condition that $a \leqslant \lambda/10$ requires that all the drivers in the 5×5 array be inside the first positive loading region in Fig. 1(a), since $(5\times2a)/2 = 5a \leqslant 0.5\lambda$. In other words, this condition simply represents the requirement for the size of the unit driver when all the drivers lie inside the first positive loading region shown in Fig. 1(a); hence, the maximum value should be achieved when the drivers are in a compact arrangement (i.e., $d=2a$). This condition is therefore a simple adaptation of the condition for an array of two drivers to the case of a 5×5 array of drivers. The difference is that the driver size specified by the condition decreases from $a=\lambda/6$ to $a=\lambda/10$ due to the increased number of drivers. Therefore, the first peak cannot be used to design a transducer array with increased radiation efficiency or output power, because the output power does not increase on increasing the number of drivers. Since the ordinate of the graphs in Fig. 4 represents the normalized mutual resistance, which is obtained by dividing the absolute mutual resistance by the characteristic acoustic impedance of the medium multiplied by the effective area of the drivers in the array $[R_m/(\rho c\times(5\times5)\pi a^2)]$, the absolute mutual resistance of a transducer array decreases as the radius of the drivers (a) decreases if the number of drivers is fixed, as in the cases shown in Fig. 4. If the magnitude of the self-resistance is taken into account, the radiation efficiency of a transducer array based on this condition imposed by the first peak in the graph of $d/a=2$ may decrease.

The analysis of the mutual resistance of two drivers presented in Sec. III A suggested two cases for maximizing the mutual effects between the drivers: (1) to arrange the two drivers in the closest possible proximity (i.e., such that $d/a=2$) if $a=\lambda/6$, and (2) to arrange the two drivers such that the center-to-center distance is $5\lambda/4$ ($d=5\lambda/4$) and $d/a=4$, if $a>\lambda/6$. As indicated in the previous paragraph, the condition given by the first peak in the region $d<\lambda/2$ is meaningless. Hence, when the mutual resistance of a 5×5 transducer array is considered, the minimum size of the unit driver is determined to be $a=\lambda/6$ because the effective radiation area increases with the size of the unit driver due to the fixed total number of drivers. It is easily observed in Fig. 4(a) that the mutual resistance of the 5×5 transducer array has its maximum when $d=5\lambda/4$ and $d/a=4$, rather than when $d/a=2$ and $a=\lambda/6$.

Figure 4(b) shows the dependence of the mutual resistance on the size of the unit driver, a/λ, for 5×5 transducer arrays characterized by a range of center-to-center distances [the corresponding plot for the two drivers system is shown in Fig. 2(b)]. It is interesting that the envelope curve of maximum mutual resistance for the 5×5 transducer array has a similar shape to the envelope observed for a system of two drivers. However, since only the region $a>\lambda/6$ is considered here, the maximum value is attained when a is approximately $5\lambda/8$ and $d/a=4$.

The graphs of mutual resistance for the 5×5 transducer arrays shown in Fig. 4 differ in several respects from those for the two driver systems in Fig. 2. The zero just before the second peak in the mutual resistance occurs at $d=\lambda$ in the case of two drivers, but at around 1.1λ in the case of 25 drivers. In general, the value of d at which this zero appears approaches the value of d for the second peak as the number of drivers increases.

B. Total radiation resistance of a hexagonal array

As described in Sec. IV A, the mutual effect in a hexagonal array is maximized when the driver radius (a) is between $\alpha\lambda$ and $\lambda/2$, and the center-to-center distance is $5\lambda/4$. Given that the total resistance of a transducer array is affected by self-resistance as well as mutual resistance, we calculated the value of the total resistance using Eq. (9). Figure 5 shows the total resistance of a hexagonal array of 25 drivers. Each curve in Fig. 5 represents the dependence of the total resistance on d/λ at constant d/a. If the size of the unit driver is fixed, then the center-to-center distances are also fixed; hence, each curve can be interpreted as representing the variation in the total resistance with changes in the driving frequency. In Fig. 5, all of the curves of total resistance for which d/a is greater than 3 show a stiff and sharp peak close to $d=5\lambda/4$, as was observed in the graphs of mutual impedance in Fig. 4(a) of Sec. IV A. In Fig. 5, several small peaks are observed in the region of $d>5\lambda/4$, corresponding to the third, fourth, and fifth peaks of the mutual impedance curves in Fig. 4(a). In particular, the magnitude and number of these peaks increase in Fig. 5 when d/a is greater than 4. For example, in all of the graphs for which d/a is greater than 6 the height of the peak close to $d=5\lambda/4$ is less than the height of the peak at $d\approx2.1\lambda$ or 2.4λ. Under the condition of constant d/a, the value of a/λ where $d=2.1\lambda$ or $d=2.4\lambda$ is two times larger than that where $d\approx5\lambda/4$, and consequently the value of the self-resistance where $d=2.1\lambda$ or 2.4λ should be larger than that where $d\approx5\lambda/4$. If d/a is less than 4, the value of a/λ at which $d\approx5\lambda/4$ is large enough to have approached the saturation value of the self-resistance; hence, the total resistance of the transducer array with $d\approx5\lambda/4$ is larger than that where d is around 2.1λ or 2.4λ.

If the drivers are in the most compact arrangement ($d/a=2$), the total resistance reaches its maximum at a relatively low frequency and then remains almost constant with some small fluctuations (see Fig. 5). However, the saturated value of the total resistance in this case is smaller than the maximum value of the total resistance for the array in which $d/a=4$. Therefore, the compact arrangement of drivers in an array ($d/a=2$) is the optimal choice for applications requiring the even production of sounds over a wide frequency band. However, to effectively produce high intensity harmonic sounds in a narrow frequency band, it is better to select a transducer array in which the drivers are arranged such that $d/a=4$.

C. Effect of the number of piston drivers

In Sec. IV A, it was shown that the mutual resistance of a transducer array with $d/a=4$ is larger than that of an array with $d/a=2$ in the narrow band neighborhood of a given frequency ($f\approx5c/4d$) when the number of drivers is sufficiently large. To investigate the effect of increasing the number of drivers in more detail, the mutual and total resistances

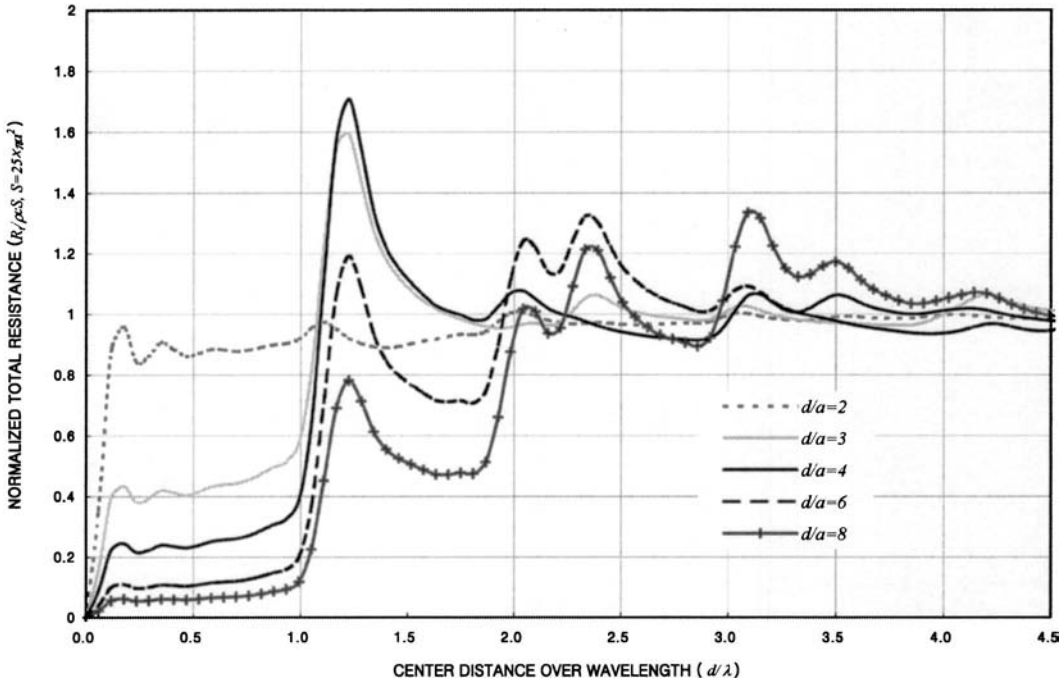


FIG. 5. Normalized total resistance of hexagonal array (5×5=25 units in an array).

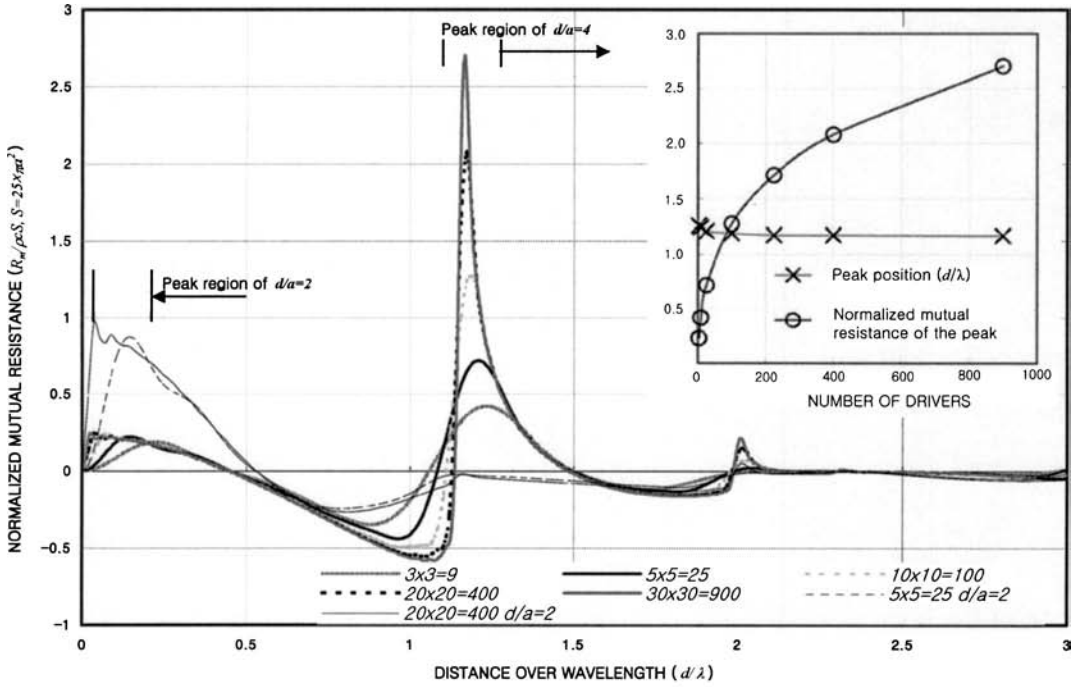
were calculated for transducer arrays ranging in size from 2×2 to 30×30. The results of these calculations are plotted in Figs. 6(a) and (b). As in Figs. 4 and 5, the mutual and total resistances are normalized by the characteristic acoustic impedance of the medium multiplied by the effective area of the transducer array [i.e., $R_t/(\rho c \times N\pi a^2)$ or $R_m/(\rho c \times N\pi a^2)$], and plotted as functions of d/λ to eliminate the effects of the increase in the radiation area caused by increasing the number of drivers.

Let us consider the mutual resistance shown in Fig. 6(a). Bold lines show the behavior of the mutual resistance for arrays with $d/a=4$ as the number of piston drivers is increased from 9(=3×3), through 25(=5×5), 100(=10×10), 400(=20×20), to 900(=30×30). The magnitude of the peak around $d=5\lambda/4$ increases with increasing number of drivers in the array. For $d<5\lambda/4$, increasing the number of drivers additionally causes the value of d at which the mutual resistance yields zero to approach the position of the peak from λ. When the number of drivers is 400, the value of d at which the peak occurs shifts from $5\lambda/4(=1.25\lambda)$ to 1.17λ. This phenomenon, in which zeroes and peaks of the mutual resistance shift closer together, is believed to be caused by the mutual effects between pairs of drivers whose center-to-center distances are greater than d, i.e., any pair of drivers that are not nearest neighbors.

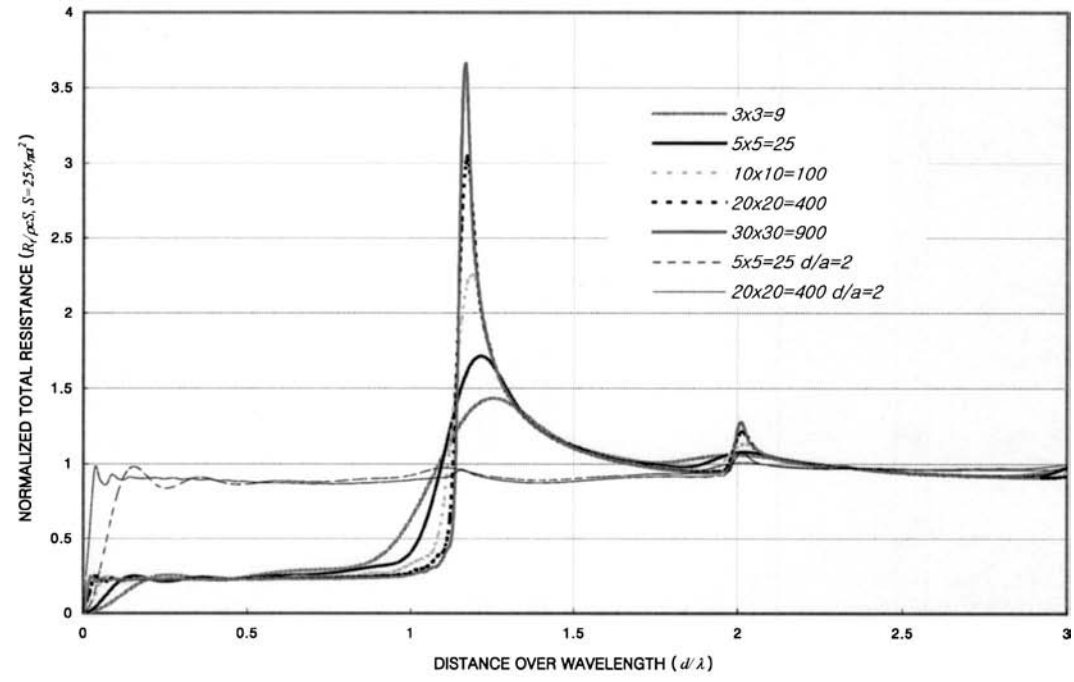
It is remarkable that the maximum value of the normalized mutual resistance (i.e., at $d\approx 5\lambda/4$) increases with increasing number of drivers. In the inset in Fig. 6(a), the values of d at which the maximum mutual resistance occurs as well as the peak values of the normalized mutual resistance are plotted as functions of the number of drivers in the array. These plots clearly show that the peak value of the normalized mutual resistance grows continuously, while the value of d at which the peaks occur approaches a constant value as the number of drivers becomes large. For example, consider the case of $d/a=2$, which is potentially the best arrangement for an array of drivers when the driver size is relatively small (i.e., $a/\lambda<\alpha$). The fine lines in Fig. 6(a) show the mutual resistances of transducer arrays of 400 and 900 drivers arranged such that $d/a=2$. In this case, the location of the maximum peak of the mutual resistance approaches the origin ($d/\lambda=0$) as the number of drivers increases. Since this peak simply indicates the condition that the entire transducer array lies inside the first peak in Fig. 1(a), as explained in Sec. IV A, the diameter of the transducer array should be almost equal to the wavelength of the generated acoustic wave; hence, this peak cannot be used to assist in the design of a transducer array with enhanced output power or power efficiency. Figure 6(b) shows the total resistance of transducer arrays of various sizes; the bold lines show the data from arrays in which $d/a=4$ and the fine lines show the results from arrays in which $d/a=2$. The maximum value of the normalized total resistance is always located around $d\approx 5\lambda/4$ when $d/a\approx 4$.

D. Application to transducer array design

We now return to the design problem outlined in Sec. III A concerning how to arrange transducers such that the resulting array gives the maximum possible radiation efficiency and output power. Assume that at a given operating frequency a unit driver can be constructed with a radius larger than $\lambda/8$ and smaller than $\lambda/2$ without any difficulty in achieving sufficient transducer internal power efficiency. However, if it is difficult to construct a unit driver of radius much larger than $\lambda/2$ with sufficiently large internal power efficiency, an easy alternative is to construct a transducer array of unit drivers of radius smaller than $\lambda/2$ in order to radiate sounds of the required output power. On the basis of the results presented here, two methods can be suggested for



(a)



(b)

FIG. 6. Effects of number of drivers to the radiation resistance ($d/a=4$, $d/a=2$): (a) effects to the normalized mutual resistance, (b) effect to the normalized total resistance.

the fabrication such a transducer array by maximizing the mutual effects of the drivers. One approach, referred to as Method I, is to construct a hexagonal-patterned transducer array in which unit drivers of radius $d/4$ are arranged such that the center-to-center distance between neighboring drivers is approximately $5\lambda/4$ (i.e., $d=5\lambda/4$ and $d/a=4$), as illustrated in Fig. 3(b). The other approach, referred to as Method II, is to construct a hexagonal-patterned transducer array in which unit drivers of radius approximately $\lambda/6$ are arranged as compactly as possible (i.e., $d/a=2$). Under the condition that the total effective radiation area of the final transducer array is maintained constant, the data shown in Fig. 6(b) clearly indicate that Method I should be superior to Method II. For example, an array of 25 drivers arranged

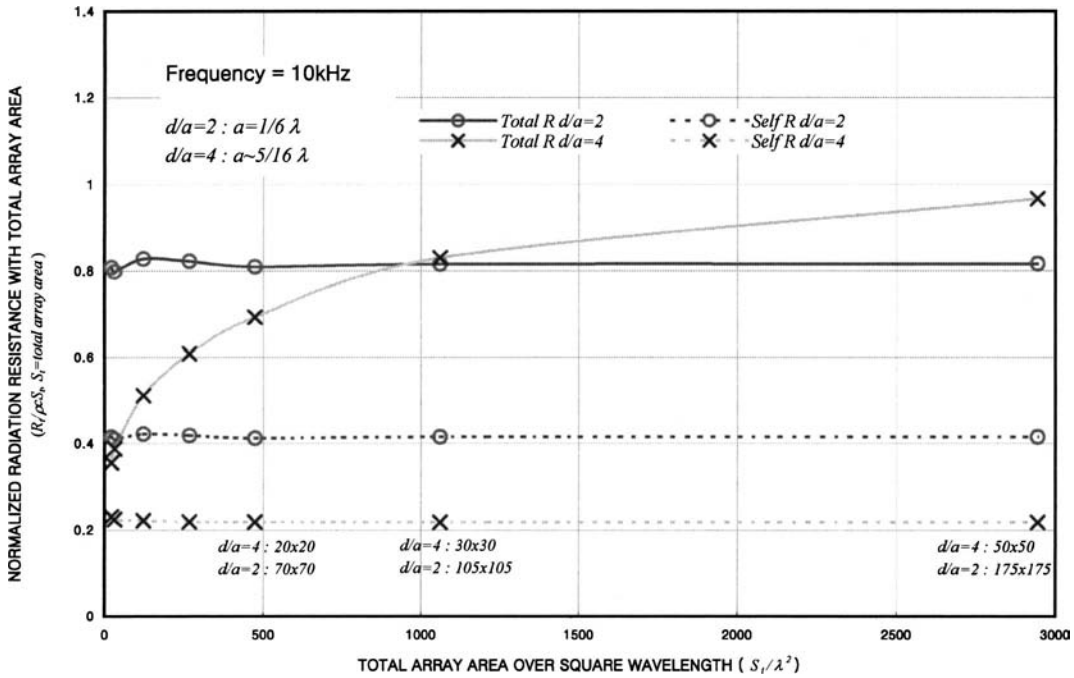


FIG. 7. The mutual radiation resistance and self-radiation resistance normalized by total array area.

using Method I achieves radiation efficiency more than 50% larger than that attained using Method II. Moreover, as more drivers are used to increase the output power, the radiation efficiency of arrays constructed according to Method I becomes higher, whereas no improvement in radiation efficiency is observed for arrays constructed using Method II. Therefore, Method I should be applied to create transducer arrays with high radiation efficiency.

In some applications the output radiation power itself is more important than the power efficiency from a given size of radiator. In such cases, the total area of the transducer array, including the effective radiating area and the area between the drivers, is fixed to a constant value. Then, the output radiation power per unit area becomes the parameter that is used to determine the output power. Thus, in this case a transducer array with larger output radiation power per unit area but lower power efficiency becomes a better choice than a transducer array with lower output radiation power per unit area but with higher radiation power efficiency. If a transducer array is constructed in a close-packed arrangement by Method II (i.e., such that $d/a=2$) with drivers of radius approximately λ/6, its total acoustic radiation area becomes larger than that of a transducer array constructed using Method I, and consequently its output power per unit area can be larger than the output power of a transducer array constructed using Method I. Thus, for a transducer array of a given total area, it is worth determining which method (I or II) gives the greatest output radiation power. To investigate this, Eq. (9) was used to calculate the total radiation resistance of transducer arrays constructed by Methods I and II for arrays of various sizes; the results are plotted in Fig. 7. The values of the radiation resistance of an array at a given frequency cited in Fig. 7 have been normalized with the characteristic acoustic impedance of the medium multiplied by the total area of the transducer array. As expected, a transducer array constructed according to Method II provides a larger output radiation power per unit area than that constructed using Method I if the total area of the transducer array is constrained such that only a relatively small number of drivers can be used. However, if the total area of the transducer array is sufficient to allow more than 900(=30 ×30) drivers in the array, the output radiation power per unit area becomes higher for a transducer array constructed by Method I, as can be seen in Fig. 7. The area in which 900 drivers can be arranged by Method I can hold approximately 11 025(=105×105) drivers when the drivers are arranged by Method II; hence, the power efficiency of the entire transducer array should be significantly lower when it is constructed by Method II. Figure 7 shows a comparison of the total radiation resistances of transducer arrays constructed using Methods I and II with various values of the total area of the transducer array. As the area becomes larger, the difference between the total resistances of the transducer arrays constructed by Methods I and II widens.

The above-outlined comparisons are more theoretical than practical, because they do not take into account the engineering problems encountered in the processes used to construct transducer arrays. However, these comparisons do make it clear that the compactly arranged transducer array is not always the best choice. The radiation efficiency is better when the transducer array is constructed from more than 25(=5×5) drivers according to Method I, i.e., when the drivers are relatively sparsely arranged. Moreover, a larger output radiation power can be achieved with a lower number of drivers and smaller effective radiation area when the transducer array is constructed using Method I. These statements are true only if the transducer array is operated at a specific frequency.

In order to investigate other radiation characteristics of a transducer array designed by Method I, such as directivity,

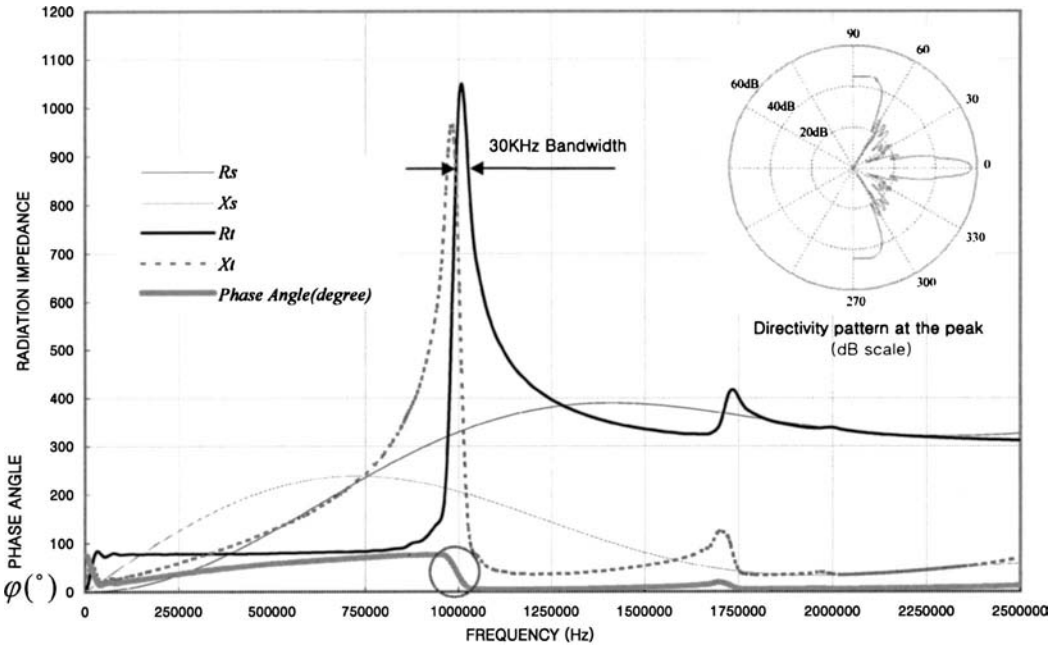


FIG. 8. The radiation impedance, phase angle, and directivity pattern at the peak range ($20\times20=400$ units in an array).

phase distortion with frequency variation, a transducer array of $400(=20\times20)$ drivers for radiating high intensity sounds of $f_c=1$ MHz in water was designed by Method I [$d/a=4$ and $d=5c/(4f_c)$], and its radiation characteristics were analyzed. The results are shown in Fig. 8. The total resistance of the transducer array (thick solid line) shows a maximum peak value at 1 MHz. The total reactance of the transducer array (thick dotted line) also has a sharp peak near the peak of the resistance. The phase of the total radiation impedance (thick gray line) varies relatively rapidly in the frequency region near the maximum peak of the total resistance ($9.72\,\text{MHz}<f_c<10.26\,\text{MHz}$), as indicated by a circle in Fig. 8. In addition, the directional characteristic of the radiated sounds (inset in Fig. 8) is shown for the frequency at which the maximum peak value of the total resistance is achieved.[12] One major lobe and one second major lobe are observed at the directions 0 and 180°, respectively, in the plot of the directional pattern of the designed transducer array. The second major lobe in the radial direction diminishes as the number of piston drivers in the array increases.

V. CONCLUSIONS

In this study, the mutual resistance of transducer arrays was investigated to determine the optimal arrangement of drivers for a high-power transducer array operating at relatively high frequency. The total resistance of a transducer array was calculated using the concepts of mutual and self-resistance based on the formulas shown in Eqs. (9) and (10). In addition to the mutual and self-resistance, the pressure waves generated by a circular piston were considered in order to analyze and explain the characteristics of the mutual effects between two circular piston drivers. The mutual resistance can be explained by interactions between the pressure waves propagated from one piston driver and the velocity of the adjacent driver. The mutual resistance of two piston drivers was analyzed to obtain theoretical explanations to describe the mutual effects. The resulting explanations were then used to predict the conditions of the local maxima of the mutual resistance curves. On the basis of the analyses of the interactions between a driver and the acoustic pressure waves, the values of the radius of a driver and the distance between two drivers that give maximum mutual radiation resistance were calculated. The widely used hexagonal arrangement of multiple drivers was found to be the optimal geometry.

The mutual resistance of a transducer array was calculated using the formula developed by Pritchard[4] to investigate the arrangement of piston drivers into an array. It was found that for efficient high power radiation of sounds the optimum distance between piston drivers is about 5/4 of the wavelength of the sounds generated by the piston drivers, and the radius of piston drivers should be 1/4 of the distance (or about 5/16 of the wavelength), assuming that efficient drivers of that size can be manufactured. The results of the calculations of the total resistances of arrays with many drivers show that the transducer array with a center-to-center distance between drivers of $d=1.17c/f$ and a driver radius of $a=1.17c/(4f)$ at a given frequency f can also achieve a larger value of the radiation power per unit area and a better radiation efficiency than an array with $d=c/(3f)$ and $a=c/(6f)$ at the same frequency.

ACKNOWLEDGMENTS

The authors would like to thank C. Cho and J. Sung in Postech for their help in some numerical calculations and helpful discussions. We also express our appreciation to K. Son and K. Han in Postech. This work is supported partly by the project ''Brain Korea 21'' from Ministry of Education, partly by the POSCO research fund, and partly by Postech.

[1] B. K. Novikov, O. V. Rudenko, and V. I. Timoshenko, *Nonlinear Underwater Acoustics* (The American Institute of Physics, New York, 1987), pp. 143–147.
[2] M. Yoneyama, J.-ichiroh Fujimoto, Y. Kawamo, and S. Sasabe, ''The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design,'' J. Acoust. Soc. Am. **73**, 1532–1536 (1983).
[3] Technical document from American Technology Corp., ''HSS Technology White Paper'' (http://www.atcsd.com).
[4] R. L. Pritchard, ''Mutual acoustic impedance between radiators in an infinite rigid plane,'' J. Acoust. Soc. Am. **32**, 730–737 (1960).
[5] T. F. Hunter, ''Twenty years in Underwater Acoustics: Generation and reception,'' J. Acoust. Soc. Am. **51**, 1025–1040 (1972).
[6] C. Audoly, ''Some aspects of acoustic interactions in sonar transducer arrays,'' J. Acoust. Soc. Am. **89**, 1428–1433 (1991).
[7] S. Shibata, T. Koda, S. Matsumoto, and J. Yamaga, ''Coaxial circular spherical array for ultrasonic imaging,'' J. Acoust. Soc. Am. **62**, 819–824 (1977).
[8] O. Mattiat, ''Transducers of producing ultrasonic waves,'' J. Acoust. Soc. Am. **25**, 291–296 (1953).
[9] F. J. Fahy, ''Acoustic interaction between structures and fluids,'' *Handbook of Acoustics*, edited by M. J. Crocker (Wiley, New York, 1998), Chap. 10, pp. 127–144.
[10] C. L. Morfey, *Dictionary of Acoustics* (Academic, New York, 2001), p. 304.
[11] D. L. Carson, ''Diagnosis and cure of erratic velocity distribution in sonar projector arrays,'' J. Acoust. Soc. Am. **34**, 1191–1196 (1962).
[12] L. Kinsler, A. Frey, A. Coppens, and J. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982), pp. 171–204.
[13] C. L. Scandrett, J. L. Day, and S. R. Baker, ''A modal Pritchard approximation for computing array element mutual impedance,'' J. Acoust. Soc. Am. **109**, 2715–2729 (2001).
[14] P. R. Stepanishen, ''An approach to computing time-dependent interaction forces and mutual radiation impedances between pistons in a rigid planar baffle,'' J. Acoust. Soc. Am. **49**, 238–292 (1970).
[15] F. Fahy, *Sound and Structural Vibration: Radiation, Transmission and Response* (Academic, New York, 1985), pp. 130–136.
[16] R. T. Richards, J. B. Bolttman III, and B. McTaggart, ''Physics of array element interaction phenomena,'' *Proceedings of the International Workshop on Power Transducers for Sonics and Ultrasonics* (Springer, Berlin, 1990), pp. 86–108.

Interlaboratory comparison of microphone calibration

George S. K. Wong and Lixue Wu
Acoustical Standards, Institute for National Measurement Standards, National Research Council Canada, Ottawa, Ontario, K1A 0R6, Canada

(Received 25 September 2003; revised 15 October 2003; accepted 15 October 2003)

The results of an interlaboratory comparison on microphone calibration involving five National Metrology Institutes (NMIs) are presented. The Institute for National Measurement Standards (INMS—Canada), was the pilot laboratory that provided the data analysis for this investigation. The final results presented here were discussed and agreed upon by all participants. The maximum rms deviation for the two LS1P laboratory standard microphones measured by the above participants is 0.037 dB, which may be considered as the key comparison reference value.
[DOI: 10.1121/1.1632091]

PACS numbers: 43.38.Kb, 43.58−e [AJZ] Pages: 680–682

I. INTRODUCTION

The Sistema Interamericano de Metrologia (SIM) interlaboratory microphone comparison designated as SIM AUV.A-K1, involved the Institute for National Measurement Standards (INMS—Canada), National Institute of Standards and Technology (NIST—U.S.A.), Centro Nacional de Metrología (CENAM—Mexico), Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO—Brazil), and Instituo Nacional de Technología Industrial (INTI—Argentina). The comparison was organized in a ''Star'' fashion, i.e., the two LS1P laboratory standard condenser microphone transfer standards were calibrated by the pilot laboratory before delivery to and after the return from a participating laboratory. The transfer dates of the microphones and the corresponding pilot laboratory measurement data had been published.[1]

II. CALIBRATION ARRANGEMENTS AT THE PILOT LABORATORY

The laboratory standard microphones (Brüel & Kjær 4160 microphones, s/n 907 045 and s/n 1 734 004) were calibrated with the reciprocity method[2] inside an environmental controlled chamber,[3] at the IEC reference conditions of 23 °C, 101.325 kPa, and 50% RH. The expanded uncertainty ($k=2$) of the temperature, pressure, and humidity measurements were 0.05 °C, 20 Pa, and 5% RH, respectively. When the microphones were returned by the participants, they were allowed to stabilize under laboratory conditions for 2 days before measurements. The microphones were then immersed in the above reference environment for at least 12 h with ISO standard air[4] before the commencement of data logging. The estimated[1] stability of the transfer standard microphones including the repeatability of the calibration system at the pilot laboratory was approximately 0.01 dB.

III. PROTOCOL AGREEMENT

The participants of the comparison were requested to perform reciprocity calibration of the above two transfer standard microphones at the following seven frequencies: 125, 250, 500, 1000, 2000, 4000, and 8000 Hz. The report from the participants was specified to include the open-circuit sensitivities of the microphones corrected to the reference environmental conditions if so desired, and with estimated expanded uncertainties ($k=2$).

IV. DATA PRESENTATION

The expanded uncertainties ($k=2$) declared by the laboratories are summarized in Table I. Figures 1 and 2 represent the measurement deviation of each laboratory from the mean of the measured values of the five laboratories at each frequency over the specified frequency range. The data submitted by the five laboratories is listed in Table II. The maximum departure of the measured values shown in the above figures is within the estimated uncertainty of 0.05 dB for reciprocity calibration of LS1P laboratory standard microphones given in IEC 1094-2 (1992).

The measurement deviations from the mean values for all laboratories shown in Figs. 1 and 2 for the two microphones were used to obtain the standard deviation for each microphone at each frequency shown in Fig. 3. It may be concluded that microphone s/n 1 734 004 exhibits relatively better stability characteristics.

TABLE I. Measurement expanded uncertainties in dB ($k=2$) quoted by each laboratory.

Freq. (Hz)	125	250	500	1000	2000	4000	8000
NRC	0.05	0.02	0.04	0.04	0.04	0.05	0.06
CENAM	0.04	0.04	0.04	0.05	0.05	0.05	0.1
INMETRO	0.05	0.05	0.05	0.05	0.05	0.07	0.11
NIST	0.04	0.04	0.04	0.04	0.04	0.04	0.12
INTI	0.05	0.05	0.05	0.05	0.05	0.1	0.1

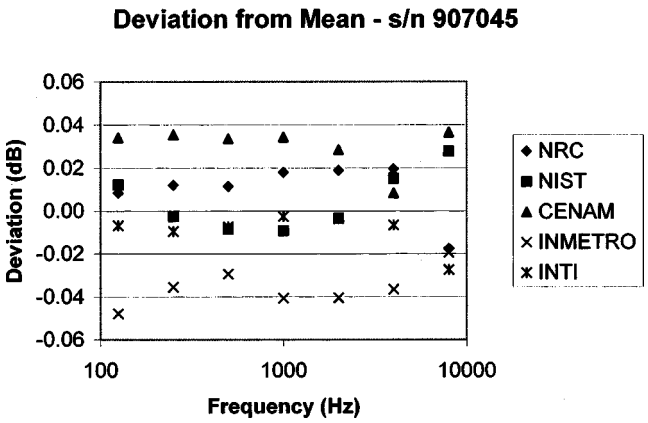


FIG. 1. Deviations of the measurement obtained by each laboratory from the mean of the measured values from all participants for microphone s/n 907 045.

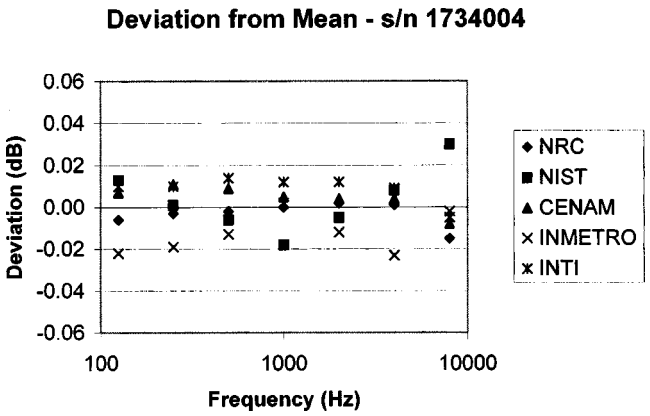


FIG. 2. Deviations of the measurement obtained by each laboratory from the mean of the measured values from all participants for microphone s/n 1 734 004.

TABLE II. Data summary.

Mic. S/N	Freq. (Hz)	NRC data[a]	NIST data[b]	CENAM data	INMETRO data	INTI data
907 045	125	−26.502	−26.500	−26.478	−26.560	−26.519
	250	−26.523	−26.537	−26.499	−26.570	−26.544
	500	−26.528	−26.549	−26.507	−26.570	−26.548
	1000	−26.491	−26.519	−26.475	−26.550	−26.512
	2000	−26.310	−26.333	−26.301	−26.370	−26.333
	4000	−25.703	−25.708	−25.715	−25.760	−25.730
	8000	−27.188	−27.143	−27.134	−27.190	−27.198
1 734 004	125	−26.988	−26.975	−26.981	−27.010	−26.978
	250	−27.009	−27.010	−27.000	−27.030	−27.001
	500	−27.016	−27.023	−27.008	−27.030	−27.003
	1000	−26.989	−27.010	−26.987	−26.990	−26.980
	2000	−26.853	−26.863	−26.854	−26.870	−26.846
	4000	−26.402	−26.399	−26.403	−26.430	−26.398
	8000	−27.350	−27.308	−27.346	−27.340	−27.343

[a]Since NRC, the pilot laboratory, measured before and after the measurements of each of the participants, the mean value (Ref. 1) of the five runs was taken as the value for NRC.
[b]The NIST data had been corrected to the reference environmental conditions by the pilot laboratory.

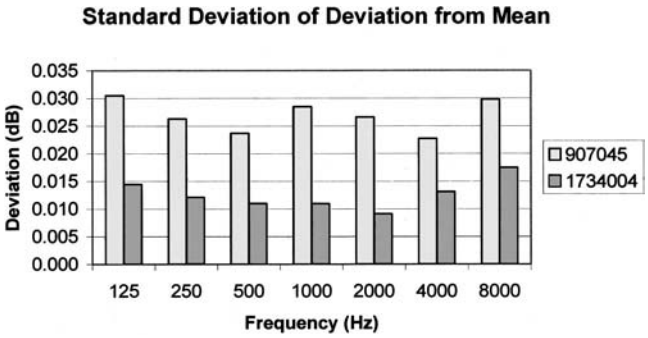


FIG. 3. For the two microphones, the deviations shown in Figs. 2 and 3 were used to compute the standard deviation from the mean of the measured values by the participants. It may be concluded that microphone s/n 1 734 004 exhibits relatively better stability characteristics.

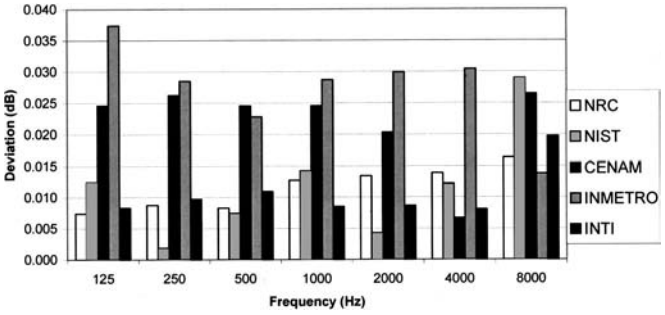


FIG. 4. The deviations of both microphones shown in Figs. 2 and 3 were used to obtain the rms deviation from the mean of the measured values for each laboratory.

To demonstrate the measurement capabilities of the participants, the measurement deviation from the mean of both microphones shown in Figs. 1 and 2 were used to obtain the rms deviation from the mean measurement for each laboratory at each frequency shown in Fig. 4. The maximum rms deviation for the two microphones measured by the NMIs is 0.037 dB, which may be considered as the key comparison reference value.

In general, all the participants used the plane-wave[2] coupler that has a volume of approximately 3 cc filled with air, except NIST, which used a large-volume coupler with a volume of approximately 20 cc filled with air for frequencies up to 500 Hz and filled with hydrogen at higher frequencies.

V. CONCLUSIONS

For the above comparison, there were five participating NMIs: the Institute for National Measurement Standards (INMS—Canada), National Institute of Standards and Technology (NIST—U.S.A.), Centro Nacional de Metrología (CENAM—Mexico), Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO—Brazil), and Instituo Nacional de Technología Industrial (INTI—Argentina). Each of the laboratories used the apparatuses of their choice, with no restriction on coupler sizes and the inclusion of gas other than air where necessary. The measurements were made at various environmental conditions with the data corrected to the IEC reference environmental conditions of 23 °C, 101.325 kPa, and 50% RH. The maximum rms deviation of microphone sensitivities from the mean value measured by the above NMIs for both microphones is 0.037 dB, which may be considered as the key comparison reference value. The overall data confirmed that the above comparisons were excellent, since IEC 1094-2 (1992) gives an estimated uncertainty of 0.05 dB for reciprocity calibration of LS1P laboratory standard microphones.

ACKNOWLEDGMENTS

The authors would like to thank the following contributors: Alfredo Elias-Juarez and Salvador Barrera-Figueroa of CENAM, Mexico; Walter E. Hoffmann and Gustavo Ripper of INMETRO, Brazil; Lucia N. Taibo and Jorge M. Riganti of INTI, Argentina, and Victor Nedzelnitsky of NIST. They all participated in the above SIM AUV.A-K1 microphone interlaboratory comparison project. Their help in supplying the measurement data is very much appreciated.

[1] G. S. K. Wong and L. Wu, "Primary microphone calibration system stability," J. Acoust. Soc. Am. **114**(2), 577–579 (2003).
[2] IEC 1094-2 Ed. 1, *Measurement microphones Part 2: Primary method for pressure calibration of laboratory standard microphones by the reciprocity technique* (International Electrotechnical Commission, Geneva, 1992).
[3] G. S. K. Wong and L. Wu, "Controlled environment for reciprocity calibration of laboratory standard microphone and measurement of sensitivity pressure correction," Metrologia **36**(4), 275–280 (1999).
[4] ISO 2533, *Standard Atmosphere* (International Organization for Standardization, Geneva, 1975).

Transient energy exchange between a primary structure and a set of oscillators: Return time and apparent damping

Antonio Carcaterra
Università Degli Studi di Roma "La Sapienza," Dipartimento di Meccanica e Aeronautica, Via Eudossiana 18, 00184 Roma, Italy

Adnan Akay[a)]
Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

(Received 5 February 2003; accepted for publication 17 November 2003)

In this paper we examine the conditions that influence the return time, the time it takes before energy returns from a set of satellite oscillators attached to a primary structure. Two methods are presented to estimate the return time. One estimate is based on an analysis of the reaction force on a rigid base by a finite number of oscillators as compared with an infinite number of continuously distributed oscillators. The result gives a lower-bound estimate for the return time. A more accurate estimation results from considering the dynamic behavior of a set of oscillators as waves in a waveguide. Such an analogy explains energy flow between a primary structure and the oscillators in terms of pseudowaves and shows that a nonlinear frequency distribution of the oscillators leads to pseudodispersive waves. The resulting approximate expressions show the influence of the natural frequency distribution within the set of oscillators, and of their number, on the return time as compared with the asymptotic case of a continuous set with infinite oscillators. In the paper we also introduce a new method based on a Hilbert envelope to estimate the apparent damping loss factor of the primary structure during the return time considering transient energy flow from the primary structure before any energy reflects back from the attached oscillators. The expressions developed for return time and damping factor show close agreement with direct numerical simulations. The paper concludes with a discussion of the return time and its relation to apparent damping and optimum frequency distribution within a set of oscillators that maximize these quantities. [DOI: 10.1121/1.1642619]

PACS numbers: 43.40.At, 43.40.Kd [RLW] Pages: 683–696

I. INTRODUCTION

In a complex system that consists of a primary, or master, structure with a finite set of oscillators attached to it, in the absence of any source of classical dissipation, the complete and permanent transfer of energy from the master to the attached oscillators amounts to an irreversible energy transfer. Irreversibility in physical systems that have finite dimensions normally develops as a result of nonlinearities in the system. For instance, in many-body systems connected with strongly nonlinear stiffness but without any source of dissipation or damping, in the classical sense, an initial energy applied to a particular mode of the system is expected to equipartition itself to all the other modes and remain there. When such a system is forced externally, it continues to absorb energy indefinitely. An example of such a physical phenomenon is friction, or the passage of a wave in a material, where atoms are excited to a higher-energy level, thus representing ''heat.'' In such cases, the energy introduced to the attached set of oscillators never reorganizes itself to return to the source, the master—giving rise to their ''thermalization.''[1] The conditions that lead to the thermalization of attached satellite oscillators develop an apparent damping of energy in the primary structure. However, linear conservative systems, the subject of this paper, behave significantly differently.

In a corresponding conservative linear system, the energy exchange between a primary structure and its satellites normally takes place with some recurrence determined by the system configuration. Considering transient cases, when a primary structure responds to an initial excitation, energy transferred to the attached oscillators returns to the primary structure after a delay during which the oscillators undergo their own oscillatory motions. Depending on the distribution of the oscillator parameters, energy may remain in the oscillators for a time, thus making them behave as a source of apparent damping for the main structure. The period during which energy remains within the oscillators may become very large.[1] However, conditions that lead to such an irreversible energy transfer in conservative linear finite systems are rare, unlike in certain nonlinear systems.

A number of investigators, in part, inspired by the fuzzy damping concept introduced by Soize,[2] examined many aspects of energy exchange between a so-called master structure, that in the simplest case is represented as a single-degree-of-freedom oscillator, and a set of oscillators attached to it (viz., Refs. 3–15). These studies differ from the earlier ones, as cited in Ref. 13, that considered the influence of internal degrees of freedom on a structure but with fewer

a)Electronic mail: akay@cmu.edu

elements, except for Ref. 16, which considered a problem similar to that described here.

Pierce *et al.*,[3] Strasberg and Feit,[4] Strasberg,[5] and others (viz., Ref. 16) showed that numerous oscillators attached to a master structure collectively act like a viscous damper and increase its effective damping significantly. Their results further suggest that damping induced by the attached oscillators on the master structure do not depend on the details of the loss factor in each oscillator and that such a set of oscillators absorbs energy, even for vanishing values of a loss factor in each oscillator, as long as the number of oscillators remain large, approaching infinity.[3]

Weaver, in a series of investigations,[7–11] generally corroborated these results, providing alternative approaches and results, and validated the expression developed by Pierce *et al.*[3] for vanishing loss factors in the oscillators for early times, pointing out that, in case of a finite number of oscillators, the assertions for vanishing loss factors hold true only for the transient exchange of energy. At later times, energy returns to the master,[7] behaving as a series of exponentially decaying pulses in the impulse response.

This subtle but significant difference relates to the issues surrounding modal overlap and the conditions under which it is permissible to represent the summation in the equations of motion by an integration, as discussed previously by, for example, Strasberg[5] and Xu and Igusa,[16] and presented unambiguously by Maidanik.[15] At a more fundamental level, however, the issue centers around the definitions of ''apparent damping'' that a master experiences as energy flows to the attached oscillators versus true dissipation that converts vibratory energy into thermal energy, through loss mechanisms in the attached oscillators. (The one exception that crosses over these two cases would be a nonlinear system, such as a large number of atoms in a solid that undergoes oscillations, which themselves represent thermal energy or true dissipation.[1,5])

Consequently, as Maidanik[15] points out, the actual dissipation of energy in a linear system requires a physical loss mechanism for the entire system. For a large number of oscillators, the loss factor in each can be small as long as their sum is finite.[15] On the other hand, if the attached oscillators have zero loss, even for a very large number of oscillators, their presence offers an apparent damping source. Drexel and Ginsberg[17] show that a cantilever beam presents an exception because the attached fuzzy structure does not absorb energy due to spatial variability unless it has heavy damping. As discussed by Weaver[7] and shown later in this paper, the apparent damping in linear systems is a transient event followed by a periodic return of energy with a return time that depends on the number of oscillators. For a very large number of oscillators, the transient event has a very long duration. The consideration of energy exchange between the master and attached oscillators in terms of a wave analogy suggests that having an infinite number of oscillators is analogous to a semi-infinite continuum where the waves do not reflect, unlike a finite-sized continuum that would be analogous to a finite number of attached oscillators. The larger the number of oscillators, the larger the continuum would be, and the longer the time for reflections. The presence of damping in the case of a semi-infinite continuum does not have any influence on the energy transmission away from the application point since there is no reflection. In the case of a finite continuum, the reflection may become very weak if the continuum has a large size or has damping. The combination of size and the presence of damping determines the reflections. Without any damping, a finite linear continuum always reflects the energy.

As previous studies have shown, achieving irreversible energy transfer in conservative linear systems requires infinite degrees of freedom. For practical cases where the primary structure has a finite number of oscillators attached to it, the concept of irreversibility applies only during a transient period described as the return time t^* during which energy flows into the satellite oscillators before returning to the primary structure, which is the subject of this paper.

In this paper we examine the conditions that influence the length of t^*, the time it takes before energy returns from a set of satellite oscillators, here called a layer, attached to it, and also introduces a method based on a Hilbert envelope to estimate the apparent damping loss factor of the primary structure during t^*.

Two different techniques are developed to produce approximate expressions for the return time. The first, described in Sec. II, calculates the reaction force of a layer of finite oscillators on a rigid base that represents the primary structure and compares it with the reaction force from an infinite set and, using asymptotic Kelvin expansion, produces an approximate expression for the return time as a function of the number of oscillators and distribution of their properties within the layer. The result is a conservative estimation for the return time. The second method, presented in Sec. III, treats the layer as a waveguide, examining the dynamic behavior of the layer in terms of waves, to obtain an expression for the return time t^*, accounting for the interaction of the primary structure and attached oscillators. The wave analogy leads to a more accurate estimate of return time and demonstrates the existence of a behavior analogous to dispersive and nondispersive waves in waveguides. The waveguide analogy also helps explain the underlying physics of the energy trapping process in the layer in terms of an interaction of two types of ''pseudowaves;'' those associated with the freely vibrating oscillators and those associated with their response when forced by the primary structure.

Simulations show how energy distribution takes place within a layer of oscillators and demonstrates the existence of a relationship between natural frequency distribution and those parts of the layer that absorb energy from the master. The paper concludes with a discussion of the relationship between the return time and apparent damping and presents expressions for the frequency distribution within a set of oscillators that maximize the return time and apparent damping, respectively.

II. REACTION FORCE AND RETURN TIME

The model considered here consists of a rigid oscillator as the primary structure and a set of oscillators individually attached to it referred here as a layer of oscillators, as sketched in Fig. 1. When the primary structure is given an

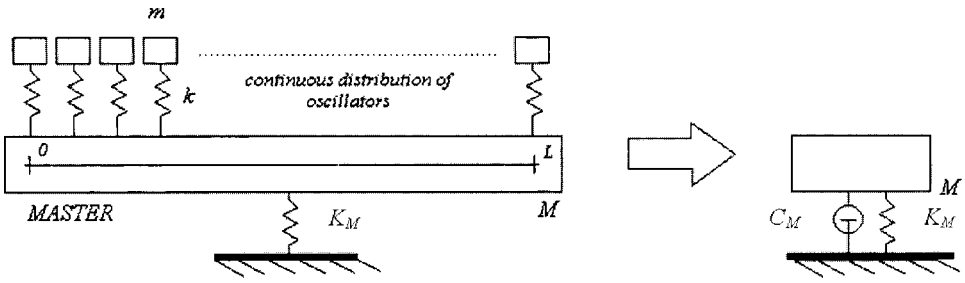


FIG. 1. Schematic description of a primary structure and a set of oscillators attached to it and representation of the layer as a viscous damper.

impulse, part or all of its energy flows to the layer. Under normal conditions, energy flows back to the master structure after a return time t^*, and thus begins an energy exchange process back and forth between the structure and the oscillators. The subject of this paper concerns the behavior of the combined system prior to the return time.

The analysis that describes the behavior of the system begins with a set of general equations of motion for a primary structure and an attached layer of continuously distributed oscillators:

$$M\ddot{x}_M + K_M x_M + \int_0^1 k(\xi)[x_M - x(\xi)]d\xi = 0, \tag{1}$$

$$m\ddot{x}(\xi) + k(\xi)[x(\xi) - x_M] = 0, \tag{2}$$

with initial conditions $x_M = 0$, $x(\xi) = 0$, $\dot{x}_M = V_0$, $\dot{x}(\xi) = 0$; $k(\xi)$ represents the stiffness distribution among the layer of oscillators, $m(\xi) = m$ is the mass distribution taken as constant in the ensuing calculations, $x_M(t)$ is the master's displacement, and $x(\xi,t)$ is the displacement distribution along the layer. [For brevity, the time variable t is omitted in $x_M(t)$ and $x(\xi,t)$.] The variable $\xi \in [0,1]$ represents the distribution domain and is dimensionless.

A typical response of the primary structure, obtained by solving the coupled equations (1) and (2) and shown in Fig. 2, shows that the layer acts as an energy sink up to a time t^* after which energy returns to the primary structure. Since

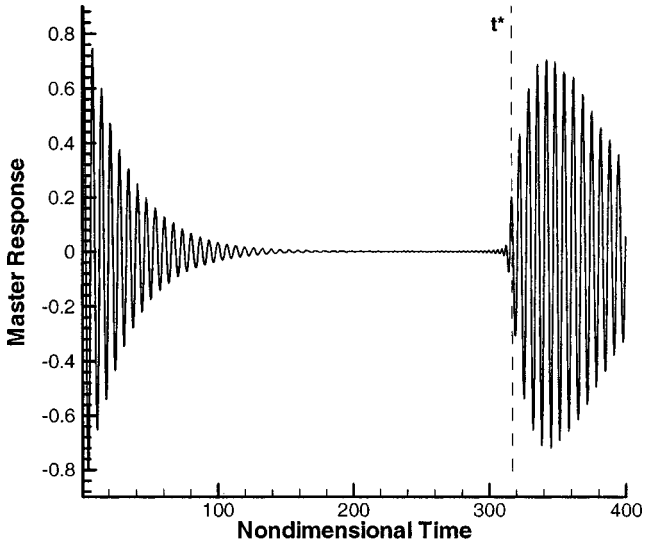


FIG. 2. Displacement response, x_M, of the primary structure in Fig. 1 to an impulse excitation as a function of nondimensional time, $t\sqrt{K_M/M}$. The layer has $N = 100$ oscillators that have a linear frequency distribution and a mass ratio of $m/M = 0.1$.

layers with small masses do not appreciably change the natural frequency of the primary structure, the transfer of energy suggests that in Eq. (1) the effect of the layer on the primary structure can be modeled with an "equivalent" viscous damper[3–5,7,16] of constant C_M up to time t^*, as described in Fig. 1:

$$M\ddot{x}_M + K_M x_M + C_M \dot{x}_M = 0, \quad t \leqslant t^*. \tag{3}$$

Thus, the term $\int_0^1 k(\xi)[x_M - x(\xi)]d\xi$ is replaced by $C_M \dot{x}_M$, where C_M is unknown and determined invoking the conservation of energy, as shown in Sec. IV. A theoretical basis for this substitution is provided later by the introduction of a suitable wave analogy.

The following analyses provide the two parameters of interest in this paper, the return time t^* and the damping coefficient C_M, and discuss the implications of having a continuous versus discrete set of oscillators.

A. Reaction by a layer of continuous oscillators

In this section we present the derivation of an expression for the total reaction force by a continuous set of parallel oscillators that act on a rigid platform.

The set of oscillators has a continuous stiffness distribution $k(\xi)$ and a constant mass distribution $m(\xi) = m$. Considering each differential mass $m\,d\xi$ as an oscillator with $k(\xi)d\xi$ as its stiffness, $\omega_n(\xi) = \sqrt{k(\xi)/m}$ represents the natural frequency of the layer at ξ.

Upon the application of a unit impulse $I_0 = 1$ on the layer, a fraction $dI_0 = I_0\,d\xi = d\xi$ acts on each differential portion $d\xi$ of the foundation, generating a force on it as

$$dF(\xi,t) = k(\xi)h(\xi,t)d\xi, \tag{4}$$

where the response to the impulse dI_0 of the differential mass at ξ is given as

$$h(\xi,t) = \frac{dI_0}{(m\,d\xi)\omega_n(\xi)} \sin[\omega_n(\xi)t]$$

$$= \frac{1}{m\omega_n(\xi)} \sin[\omega_n(\xi)t]. \tag{5}$$

The total reaction force on the rigid platform by the set in response to an impulsive excitation is an integral of the reaction force from each element:

$$F_C(t) = \int_0^1 k(\xi)h(\xi,t)d\xi = \int_0^1 \omega_n(\xi)\sin[\omega_n(\xi)t]d\xi. \tag{6}$$

Similar expressions for force in the remainder of this paper implicitly assume the response to a unit impulse.

For large values of time, the value of the integral in Eq. (6) can be estimated using asymptotic expansion for two possible conditions. In the absence of a stationary point in the distribution of $\omega_n(\xi)$, i.e., $\omega_n'(\xi)\neq 0$, $\forall \xi\in[0,1]$,

$$F_C(t)=\frac{1}{t}\left[\frac{\omega_n(0)}{\omega_n'(0)}\cos[\omega_n(0)t]-\frac{\omega_n(1)}{\omega_n'(1)}\cos[\omega_n(1)t]\right]. \tag{7}$$

In cases where a stationary point exists at ξ_0, i.e., $\omega_n'(\xi_0)=0$, $\xi_0\in[0,1]$,

$$F_C(t)\approx\sqrt{\frac{2\pi}{t|\omega_n''(\xi_0)|}}\,\omega_n(\xi_0)\sin[\omega_n(\xi_0)t+\theta], \tag{8}$$

where $\theta=(\pi/4)\mathrm{sgn}(\omega_n''(\xi_0))$ is obtained by using the Kelvin stationary-phase method.

The presence of a stationary point in the frequency distribution changes the time dependence of the layer reaction force from $1/t$ to $1/\sqrt{t}$. These results suggest that a set of continuously distributed oscillators, uniformly excited by an impulse, provide a total force on their common, rigid base that tends to zero as $1/t$ or $\sqrt{1/t}$, regardless of the natural frequency distribution within the layer. The result that the force $F(t)$ decreases with increasing time implies that the coherence of the oscillator vibration progressively vanishes and, in turn, suggests that an irreversible process takes place within the continuous distribution of oscillators. As coherence vanishes among the oscillators, they never approach a constructive interaction among them to return the initial impulse to the primary structure. The combined reaction force of the layer decays with a dominant frequency that corresponds to the frequency of the stationary point or in case where a stationary point does not exist, it corresponds to the dominant frequencies $\omega_n(0)$ and $\omega_n(1)$.

Examples: Linear and exponential distributions. For a layer that has a linear distribution of natural frequencies $\omega_n(\xi)=a\xi$, where a is a constant, the exact integral of the total reaction force becomes

$$F_C(t)=\frac{1}{at^2}\int_0^1 \xi\sin(\xi t)d\xi, \tag{9}$$

which trivially leads to

$$F_C(t)=\frac{1}{at^2}\sin(at)-\frac{1}{t}\cos(at), \tag{10}$$

where after a short time the first term vanishes and the result becomes equal to the asymptotic expression obtained above. Similarly, for a layer that has an exponential natural frequency distribution $\omega_n(\xi)=ae^{\beta\xi}$, the force integral can be evaluated by a substitution of variables:

$$F_C(t)=\frac{1}{t}[\cos at-\cos ae^{\beta}t]. \tag{11}$$

The asymptotic decay of the reaction force is a consequence of the continuous representation of the layer of oscillators. As seen in the following section, however, discretization of the layer makes the decay time finite and introduces a reflection from the layer.

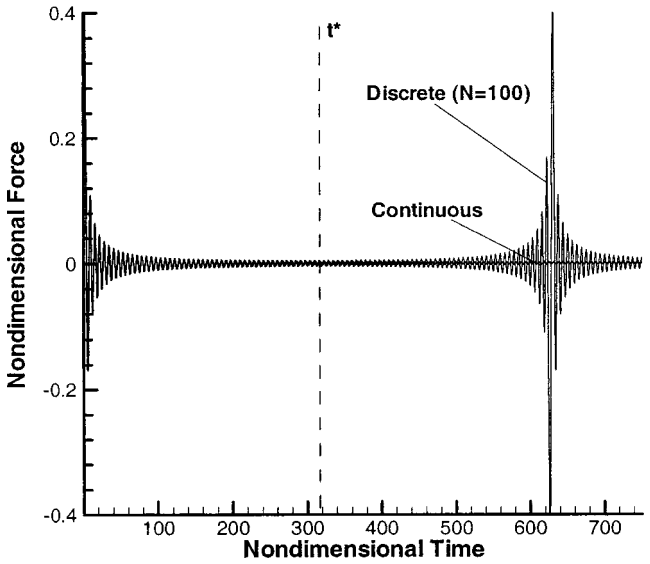


FIG. 3. A comparison of the reaction force (F/a) produced by a discrete set of $(N=100)$ oscillators given by Eq. (10) with that produced by a layer of a continuously distributed oscillators as a function of nondimensional time at.

B. Reaction by a layer of discrete oscillators

The development of the expression for reaction force $F(t)$ in the previous section assumes a continuous distribution of oscillators, implying an infinite number of them positioned along the layer. In the case of a finite number of oscillators, $F(t)$ does not vanish irreversibly, as discussed in the last section, but, as shown in Fig. 3, it develops to its original form after a time t^*.

The discrete counterpart of the force expression in Eq. (6) for a continuous distribution takes the form

$$F_D^{(N)}(t)=\sum_{i=1}^{N}\omega_n(\xi_i)\sin[\omega_n(\xi_i)t]\Delta\xi. \tag{12}$$

When the set of oscillators has a linear frequency distribution with discrete frequencies that are $\Delta\omega$ apart, Eq. (12) can be rewritten as

$$F_D^{(N)}(t)=\sum_{i=1}^{N-1}(\omega_0+i\Delta\omega)\sin[(\omega_0+i\Delta\omega)t]\Delta\xi. \tag{13}$$

The summation in Eq. (13) can be evaluated as

$$F_D^{(N)}(t)=\Delta\xi\left[-N\Delta\omega\frac{\cos[(2N-1)\Delta\omega/2]t}{2\sin(\Delta\omega/2)t}+\Delta\omega\frac{\sin(N\Delta\omega)t}{4\sin^2(\Delta\omega/2)t}\right]. \tag{14}$$

The return time for this linear set of oscillators corresponds to $t^*=2\pi/\Delta\omega$, the value that makes the denominator of Eq. (14) equal to zero. With $N\Delta\omega=\omega_{\max}$, for very large values of N, Eq. (14) reduces to the result for a continuous set of oscillators, as shown in Eq. (10):

$$\lim_{N\to\infty} F_D^{(N)}(t) = F_C(t) = \frac{1}{\omega_{\max} t^2}\sin\omega_{\max} t - \frac{1}{t}\cos\omega_{\max} t. \tag{15}$$

C. Continuous versus discrete set of oscillators

In this section we examine the differences between a continuous and discrete set of oscillators for the more complex case of a general frequency distribution and compare the force expression in Eq. (6) with its discrete counterpart to obtain the value of t^* that results from discretization.

The continuous layer over the domain $\xi\in[0,1]$ has an infinite number of differential oscillators characterized by their natural frequency spectrum $\omega_n(\xi)$. The corresponding discrete layer, on the other hand, has a finite number of N oscillators characterized by the discrete frequency spectrum $\omega_n(\xi_i)$, where $\xi_i = i/N$ also gives their positions over the same layer. For example, for a linear distribution of frequency $\omega_n(\xi) = \omega_{\max}\xi$, where $\omega_{\max}$ is selected as the upper limit of the frequency spectrum, the discrete frequencies can be expressed as

$$\omega_n(\xi_i) = (i\Delta\xi)\omega_{\max}. \tag{16}$$

The analysis starts with an equivalent discrete form of the reaction force, Eq. (6), and a residual,

$$F_C(t) = \int_0^1 \omega_n(\xi)\sin\omega_n(\xi)t\,d\xi$$

$$= \Delta\xi\sum_{i=0}^{N} \omega_n(\xi_i)\sin\omega_n(\xi_i)t + \mathcal{R}_{\mathcal{N}}(t), \tag{17}$$

where the first term (the summation) on the right-hand side represents the force applied to the foundation by a discrete set of oscillators in the layer:

$$F_C(t) = F_D^{(N)}(t) + \mathcal{R}_{\mathcal{N}}(t).$$

Following the method outlined in the Appendix, the residual $\mathcal{R}_{\mathcal{N}}(t)$ in the expanded integral (17) can be written as

$$\mathcal{R}_{\mathcal{N}}(t) = \frac{\Delta\xi}{2}\left\{\frac{d}{d\xi}[\omega_n(\xi)\sin\omega_n(\xi)t]\right\}_{\xi=\xi_N(t)}, \tag{18}$$

where $\xi_N(t)\epsilon[0,1]$. [Note that since $\xi_N(t)$ depends on the function that is integrated, a particular $\xi_N(t)$ exists for each value of t considered here.]

An upper bound for the residual $\mathcal{R}_{\mathcal{N}}(t) = F_C(t) - F_D^{(N)}(t)$ can be determined by expanding the derivative in the expression and rewriting it as (18):

$$\mathcal{R}_{\mathcal{N}}(t) = \frac{\Delta\xi}{2}\{\omega_n'(\xi)\sin\omega_n(\xi)t$$

$$+ \omega_n(\xi)\omega_n'(\xi)t\cos\omega_n(\xi)t\}_{\xi=\xi_N(t)}, \tag{19}$$

$$\mathcal{R}_{\mathcal{N}}(t) = \frac{\Delta\xi}{2}\{\mathcal{A}(\xi)\sin[\omega_n(\xi)t + \phi(\xi)]\}_{\xi=\xi_N(t)}, \tag{20}$$

where

$$\mathcal{A}(\xi) = [\omega_n'^2(\xi) + \omega_n^2(\xi)\omega_n'^2(\xi)t^2]^{1/2}, \tag{21}$$

$$\phi(\xi) = \arctan[\omega_n(\xi)t].$$

An upper limit for the residual follows as

$$|\mathcal{R}_N(t)| \leq \frac{\Delta\xi}{2}\max\{\mathcal{A}[\xi_N(t)]\}. \tag{22}$$

Although both $\mathcal{R}_{\mathcal{N}}(t)$ and $\xi_N(\xi_N\in[0,1])$ are unknown, expression (22) provides an upper-bound estimate of the difference $|F_C - F_D^{(N)}|$ that can be determined by considering the maximum value $A(\xi)$ can take:

$$\max\{\mathcal{A}[\xi_N(t)]\} \leq \max\{|\omega_n'|\}\sqrt{1+\max\{\omega_n^2\}t^2}.$$

Since $N\Delta\xi = 1$, substituting the maximum value of $\mathcal{A}$ produces the maximum value the residual can take:

$$|\mathcal{R}_N(t)| = |F_C(t) - F_D^{(N)}(t)|$$

$$\leq \frac{1}{2N}\max\{|\omega_n'|\}\sqrt{1+\max\{\omega_n^2\}t^2}, \tag{23}$$

where all the terms on the right-hand side are known quantities. This result shows that the upper bound of the difference between the forces resulting from continuous and discrete sets of oscillators depends also on time. This result shows that beyond a return time t^*, the residual $\mathcal{R}_{\mathcal{N}}$ is no longer negligible.

D. Return time t^*: An estimate for lower bound

The return time, found as $t^* = 2\pi/\omega$ for the simple case of a set of oscillators with a linear frequency distribution, becomes more difficult to predict for a set of oscillators with more complex frequency distributions

The return time also represents the time when the reaction force from a set of finite (discrete) oscillators begins to diverge from the corresponding result due to a continuous set of oscillators. A conservative estimate that gives a lower bound for t^* can be found as follows. Considering the value of the residual in Eq. (23) for large times, $\omega_{n,\max}t > 1$,

$$|\mathcal{R}_{\mathcal{N}}(t)| = |F_C(t) - F_D^{(N)}(t)| \leq \frac{1}{2N}\max\{|\omega_n'|\}\max\{\omega_n\}t. \tag{24}$$

Since $F_C(t)\to 0$ for large values of time t, expression (24) reduces to

$$|F_D^{(N)}(t)| \leq \frac{1}{2N}\max\{|\omega_n'|\}\max\{\omega_n\}t. \tag{25}$$

To simplify the analysis, the return time is taken here as the time at which the amplitude of the residual force reaches the amplitude of the first peak of the force $F_C(t)$. Thus, if $F_C(t)$ reaches its first peak value at t_1, then $F_D^{(N)}(t)$ reaches the peak value $F_C(t_1)$ at $t = t^*$, yielding an equivalent expression as

$$|F_D^{(N)}(t^*)| \sim |F_C(t_1)|. \tag{26}$$

The time when $F_C(t)$, given in Eq. (6), reaches its first peak value can be found from its derivative for small values of t (early oscillations):

$$\frac{dF_C}{dt}=\int_0^1 \omega_n^2(\xi)\cos[\omega_n(\xi)t]d\xi$$

$$\sim \int_0^1 \omega_n^2(\xi)\left[1-\frac{1}{2}\omega_n^2(\xi)t^2\right]d\xi=0,$$

which gives the time when the first peak develops as

$$t_1=\sqrt{\frac{2\int_0^1\omega_n^2(\xi)d\xi}{\int_0^1\omega_n^4(\xi)d\xi}}. \tag{27}$$

Since t_1 occurs at early times,

$$F_C(t_1)=\int_0^1 \omega_n(\xi)\sin(\omega_n t_1)d\xi \simeq t_1\int_0^1\omega_n^2(\xi)d\xi. \tag{28}$$

Substituting for t_1 from above,

$$F_C(t_1)\simeq\sqrt{\frac{2[\int_0^1\omega_n^2(\xi)d\xi]^3}{\int_0^1\omega_n^4(\xi)d\xi}}. \tag{29}$$

The equivalence of Eqs. (29) and (25) produces the return time as

$$t^*\geqslant\frac{2\sqrt{2}N}{\max\{|\omega_n'|\}\max\{\omega_n\}}\frac{[\int_0^1\omega_n^2(\xi)d\xi]^{3/2}}{[\int_0^1\omega_n^4(\xi)d\xi]^{1/2}}. \tag{30}$$

A lower bound for the return time can be obtained for any layer with a known frequency distribution. For example, for a layer of oscillators having a frequency distribution:

$$\omega_n(\xi)=\omega_{\max}\xi^p, \quad \xi\in[0,1], \tag{31}$$

yields a lower bound estimate as

$$t^*\geqslant\frac{(4p+1)^{1/2}}{(2p+1)^{3/2}}\frac{2\sqrt{2}N}{p\omega_{\max}}. \tag{32}$$

For a linear frequency distribution ($p=1$), the return time prediction reduces to

$$t^*\geqslant\sqrt{\frac{40}{27}}\frac{N}{\omega_{\max}}=\sqrt{\frac{40}{27}}\frac{1}{\Delta\omega}. \tag{33}$$

The expression (30) provides only a rough estimate for t^* giving a lower bound for the return time and also shows how t^* depends on the distribution of the natural frequencies of the oscillators in the layer. In the next section a better estimate is developed using Eq. (17) and developing a wave analogy.

III. WAVE ANALOGY AND RETURN TIME

The observations in the last section on periodic force buildup and energy return to the primary structure suggest a behavior analogous to the reflection of waves. In this section we present a wave analogy to describe the interaction between a primary structure and its oscillators with a plane-wave pulse that travels in a finite waveguide. As shown later, such a wave analogy makes it easier to estimate the return time for arbitrary, nonlinear frequency distributions.

In a waveguide, such as a rod with length L and speed of sound c, a plane-wave pulse travels between its two ends with a period $T=2L/c$. The response of such a waveguide to a unit impulse exhibits the same form as the total force from discrete set of oscillators acting on the primary structure expressed by (12):

$$\ddot{u}(0,t)=-(2/m_w)\sum_{i=1}^{\infty}\omega_i\sin(\omega_i t), \tag{34}$$

where $\ddot{u}(t)$ is the acceleration, m_w represents the mass of the waveguide, and the natural frequencies, ω_i, for the case of a rod with free ($x=0$) and fixed ($x=L$) boundary conditions are given as

$$\omega_i=\frac{\pi i c}{L}. \tag{35}$$

The frequency spacing between the natural frequencies is constant,

$$\Delta\omega=\frac{\pi c}{L}, \tag{36}$$

and leads to an expression for the speed of sound in terms of frequency spacing,

$$c=\Delta\omega\frac{L}{\pi}. \tag{37}$$

Substituting for c in $T=2L/c$ gives the time it takes for a plane wave to travel and return,

$$T=\frac{2\pi}{\Delta\omega}, \tag{38}$$

which has the same value as the periodicity of the force expression for a discrete set of oscillators given in Eq. (12).

The phase difference between the oscillators that results from the differences in their frequencies produces an analogous speed of sound for a disturbance that travels in the nondimensional ξ domain along the layer of oscillators. Such an average speed of sound represents travel along the length of the layer in a time period ΔT:

$$c_L=\frac{\xi_{\max}}{\Delta T}=\frac{1}{\Delta T}, \tag{39}$$

since $\xi_{\max}=1$. The time period during which a disturbance travels in the ξ domain ($\xi\in[0,1]$) corresponds to the phase difference between the first and the last oscillator:

$$\Delta T=T_1-T_N=\frac{2\pi}{\Delta\omega}\left(1-\frac{1}{N}\right)\approx\frac{2\pi}{\Delta\omega} \tag{40}$$

for $N\gg1$. The speed of sound c_L together with a linear frequency distribution for the oscillators as $\omega_i=\omega_{\max}(i/N)$ yields a corresponding wave number as

$$k_i=\frac{\omega_i}{c_L}=2\pi i. \tag{41}$$

It follows that the time it takes for a disturbance to reflect is

$$t^*=\frac{\xi_{\max}}{c_L}=\frac{2\pi}{\Delta\omega}, \tag{42}$$

which is the same as that obtained from the closed-form solution earlier and reported by Ref. 7.

The selected value of the maximum frequency $\omega_{\max}$ of the distribution in the layer also sets the maximum frequency of the waveguide for its equivalence with the layer. Then, for a given value of $\omega_{\max}$ for the set, the number of oscillators N corresponds to the length of the waveguide. Increasing N in the layer, corresponds to an increase in L of the waveguide. This implies that according to the present analogy, as $N \rightarrow \infty$, a perturbation must travel along an infinite waveguide without any reflection at the opposite end. In such cases, the frequency distribution tends to be continuous and the return time tends to be infinite; for the case of a linear frequency distribution, $t^* = \lim_{\Delta\omega \rightarrow 0}(2\pi/\Delta\omega)$.

The wave analogy described above explains the energy exchange process and the characteristic return time in terms of pseudowaves. The example used above for a linear frequency distribution for the oscillators can be extended to cases with more complex, nonlinear frequency distributions that lead to the development of dispersive pseudowaves.

The following two sections describe the dispersive pseudowaves and their respective group velocities. The first examines the case for the free vibrations of a layer of oscillators and the second considers the case of forced vibration of the layer by the primary structure vibrations.

A. Dispersive pseudowaves

The selection of a nonlinear frequency distribution for the natural frequencies of the oscillators produces dispersive pseudowaves with a corresponding group velocity. The wave analogy can be extended to dispersive cases by considering a general class of nonlinear frequency distributions, expressed as

$$\omega_n(\xi_i) = \omega_i = \omega_{\max}(i\,\Delta\xi)^p = \omega_{\max}\xi^p = \omega_{\max}(i/N)^p, \quad (43)$$

where p $(p>0)$ is a real constant.

The corresponding phase velocity is

$$c_\phi = \frac{\omega_i}{k_i} = \frac{\overline{\Delta\omega}}{2\pi}\xi^{p-1} = \frac{\overline{\Delta\omega}}{2\pi}\left(\frac{i}{N}\right)^{p-1}, \quad (44)$$

where $\overline{\Delta\omega} = \omega_{\max}/N$ is the average frequency spacing in the layer.

The corresponding group velocity is

$$c_g = \left.\frac{d\omega_i}{dk_i}\right|_{\omega=\bar{\omega}} = \left. p\,\frac{\overline{\Delta\omega}}{2\pi}\xi^{p-1}\right|_{\omega=\bar{\omega}} = \left. p\,\frac{\overline{\Delta\omega}}{2\pi}\left(\frac{i}{N}\right)^{p-1}\right|_{\omega=\bar{\omega}}$$
$$= p\;\; c_\phi. \quad (45)$$

With the wave analogy but for a dispersive waveguide, the return time for the set of oscillators t^* is expressed in terms of group velocity c_g as

$$t^* = \frac{\xi_{\max}}{c_g(\bar{\omega})}, \quad (46)$$

where the group velocity $c_g(\bar{\omega})$ is evaluated at $\omega = \bar{\omega}$ that represents the dominant frequency of the group waves as discussed in the development of Eqs. (7) and (8). In case of free vibrations, the dominant frequency occurs at $\xi = \xi_{\max} = 1$ or, equivalently, at $i = N$, leading to

$$c_g(\bar{\omega}) = p\,\frac{\overline{\Delta\omega}}{2\pi}. \quad (47)$$

The corresponding return time then follows as

$$t_L^* = \frac{\xi_{\max}}{c_g(\bar{\omega})} = \frac{2\pi}{p\,\overline{\Delta\omega}}. \quad (48)$$

The return time t_L^*, as in the case of nondispersive pseudowaves, represents the return time of energy from a layer of oscillators excited by an initial pulse. In a sense, this return time corresponds to the ''free vibrations'' of the uncoupled set of oscillators. The consideration of the dynamics of the primary structure, as shown below, brings in the natural frequency of the impulsively excited primary structure, setting the stage for energy exchange between the primary structure and the oscillators.

B. Pseudowaves forced by the primary structure

Because of their small total mass relative that of the primary structure, the oscillators in the layer undergo forced vibrations at the natural frequency of the primary structure following application of an impulse to it. As a result of this new dominant frequency, another set of analogous dispersive waves develops with a respective group velocity in the combined primary structure and set of oscillators. The new group velocity also results from Eq. (45) by evaluating it at $\xi = \xi_M$, where $\xi_M = (\omega_M/\omega_{\max})^{1/p}$ is the location along the layer that corresponds to the free-vibration frequency, ω_M, of the master structure. The characteristic return time associated with the forced vibrations follows from Eq. (46) as

$$t_M^* = \frac{2\pi N}{p\,\omega_M}\left(\frac{\omega_M}{\omega_{\max}}\right)^{1/p}. \quad (49)$$

The simultaneous presence of these two groups of pseudowaves has an important role in the energy exchange phenomenon that takes place between the primary structure and a set of parallel oscillators. As shown later through numerical simulations, the two return times t_L^* and t_M^* associated with the ''free vibrations'' of the layer of oscillators and the ''forced vibrations'' due to oscillations of the master, respectively, appear simultaneously. Each return time has associated with it an energy returned to the primary structure. Simulations reported later show that irrespective of which of the two arrives back at the primary structure first, most of the returned energy is associated with forced vibrations by the primary structure.

For a given mass ratio between the primary and attached oscillator set, the two return times differ according to the frequency distribution of the attached oscillators. A simple comparison between the return times in (48) and in (49) leads to the ratio

$$\frac{t_M^*}{t_L^*} = \left(\frac{\omega_M}{\omega_{\max}}\right)^{(1-p)/p}. \quad (50)$$

For effective energy exchange, the natural frequency of the primary structure always falls within the frequency distribution of the layer, i.e., $0 \leqslant \omega_M \leqslant \omega_{\max}$; thus the ratio in the parentheses in Eq. (50) never exceeds unity.

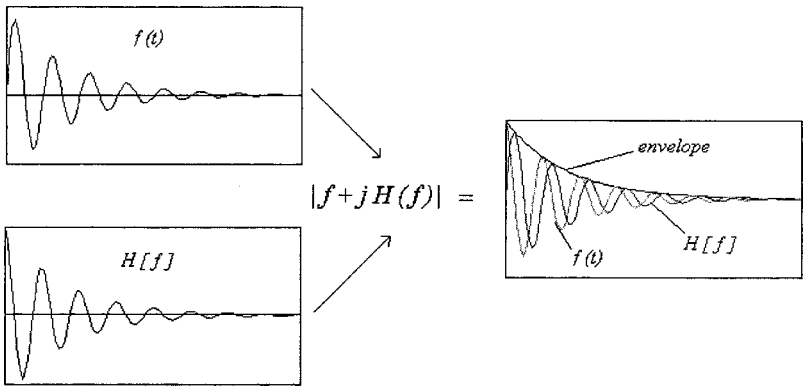


FIG. 4. The description of the Hilbert envelope of a function.

In Eq. (50), the value $p=1$ corresponds to a linear distribution of frequencies of the oscillators, producing nondispersive pseudowaves. In this special case, regardless of the ratio $(\omega_M/\omega_{\max})$, return times predicted by Eq. (48) and Eq. (49) have the same value.

The nonlinear frequency distributions that correspond to $p<1$, in Eq. (50), produce a faster pseudowave associated with the primary structure than that associated with the free vibrations of oscillators. The simulations for $p<1$ show that energy returns to the primary structure, as predicted by Eq. (48). Because of their low energy and longer return time, free-vibration pseudowaves of the oscillators become masked by the ''forced vibration'' pseudowaves.

For frequency distributions with $p>1$, delineation of the returned waves due to free vibrations of the oscillators becomes easier since they arrive earlier than those due to the primary structure, however, still with much less energy than the group waves associated with the primary structure that follow.

IV. TRANSIENT ENERGY EXCHANGE AND APPARENT DAMPING: χ

Energy flow during the return time t^* from the primary structure to the layer forms the basis of apparent damping, for the loss of energy appears as damping to the primary structure. The following analysis considers the initial system response until energy returns to the primary structure to identify apparent damping C_M in Eq. (3), where it represents the absorption of energy by the oscillators during t^*. The analysis in this section starts with the equations of motion of the system, Eqs. (1) and (2), and develops an expression for an apparent loss factor based on the Hilbert envelope.[18]

Letting $k(\xi)/m=\omega_n^2(\xi)$ and rearranging the terms in Eq. (1), $\omega_n^2(\xi)x_M(t)$ becomes the ''forcing'' term,

$$\ddot{x}(\xi,t)+\omega_n^2(\xi)x(\xi,t)=\omega_n^2(\xi)x_M(t), \tag{51}$$

with the solution

$$x(\xi,t)=\omega_n(\xi)\int_0^t \sin[\omega_n(\xi)(t-\tau)]x_M(\tau)d\tau. \tag{52}$$

The method presented here relies on the hypothesis that for large values of t, $x_M(t)\to 0$. Under these conditions, and without any classical dissipation source in the system, the layer of oscillators effectively acts as a source of damping on the motion of the primary structure. Then the integral expression (52) for the asymptotic response of the layer takes the form

$$x(\xi,t)\simeq\omega_n(\xi)\int_0^{+\infty} \sin[\omega_n(\xi)(t-\tau)]x_M(\tau)d\tau. \tag{53}$$

Since $x_M(t)=0$ for $t\leqslant 0$,

$$x(\xi,t)\simeq\omega_n(\xi)\int_{-\infty}^{+\infty} \sin[\omega_n(\xi)(t-\tau)]x_M(\tau)d\tau. \tag{54}$$

The Hilbert envelope of $x(\xi,t)$ is defined as

$$|\hat{x}(\xi,t)|=|x(\xi,t)+j\tilde{x}(\xi,t)|, \tag{55}$$

where $\tilde{x}(\xi,t)$ is the Hilbert transform of $x(\xi,t)$. As demonstrated in Fig. 4 using a decaying-sine function, the Hilbert transform acts on a signal as a phase-shift operator, and together with the function itself produces an envelope of the signal as described by Eq. (55). Applying the Hilbert transformation to the convolution integral in Eq. (54) yields

$$\tilde{x}(\xi,t)\simeq-\omega_n(\xi)\int_{-\infty}^{+\infty} \cos[\omega_n(\xi)(t-\tau)]x_M(\tau)d\tau. \tag{56}$$

Substituting the expression for $\tilde{x}(\xi,t)$ in Eq. (55) yields

$$\hat{x}(\xi,t)\simeq-\omega_n(\xi)\int_{-\infty}^{+\infty} je^{j\omega_n(\xi)(t-\tau)}x_M(\tau)d\tau, \tag{57}$$

where $\hat{x}(\xi,t)$ is the analytic signal associated with $x(\xi,t)$.

The Hilbert envelope follows as

$$\begin{aligned}|\hat{x}(\xi)|&\simeq\omega_n(\xi)\left|\int_{-\infty}^{+\infty} x_M(t)e^{-j\omega_n(\xi)t}\,dt\right|\\&=\omega_n(\xi)|\mathcal{F}(x_M)|_{\omega=\omega_n(\xi)},\end{aligned} \tag{58}$$

where $\mathcal{F}(x_M)$ represents the Fourier transform of the primary structure oscillations evaluated at the layer frequency distribution. An equivalent expression for $\mathcal{F}(x_M)$ results from applying an impulse $I=MV_0$ to the primary structure. Considering Eq. (3), which describes the response of the primary structure, equivalent expressions for its response follow as

$$\left|\int_{-\infty}^{+\infty} x_M(t)e^{-j\omega_n(\xi)t}\,dt\right|=\left|\frac{I_0}{-M\omega^2+j\omega C_M+K_M}\right|_{\omega_n(\xi)} \tag{59}$$

Thus, the asymptotic amplitude of the layer as $t\to\infty$ for any value of ξ can be represented by

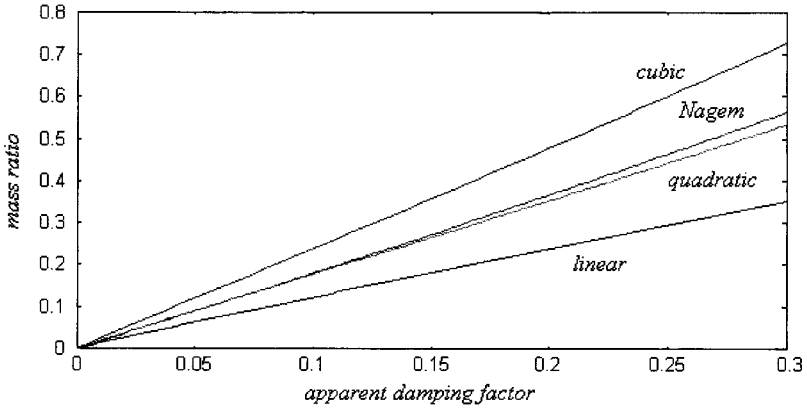


FIG. 5. Variation of the apparent damping factor with the mass ratio of layer to primary structure.

$$A(\xi)=|\hat{x}(\xi)|\simeq\omega_n(\xi)\left|\frac{I_0}{-M\omega_n^2(\xi)+j\omega_n(\xi)C_M+K_M}\right|, \tag{60}$$

where the unknown equivalent, or apparent, damping value C_M will result from the energy equation as follows.

A. Energy conservation

The total energy that the layer attains asymptotically can be expressed in an integral form as

$$E_L=\frac{1}{2}\int_0^1 k(\xi)A^2(\xi)d\xi. \tag{61}$$

Assuming that all of the initial kinetic energy induced in the primary structure $E_M=1/2MV_0^2$, transfers to the layer as $t\rightarrow\infty$ $(E_L=E_M)$,

$$\int_0^1 k(\xi)A^2(\xi)d\xi-MV_0^2=0. \tag{62}$$

Substituting Eq. (60) into (62) yields a relationship for the loss factor $\chi=C_M/M\omega_M$ in terms of mass ratio and frequency distribution:

$$\mu\int_0^1\left|\frac{\Omega_n^2(\xi)}{1-\Omega_n^2(\xi)+j\Omega_n(\xi)\chi}\right|^2 d\xi-1=0, \tag{63}$$

where the nondimensional quantities are as follows: $\mu=m_L/M$ is the mass ratio of the total layer to the primary, $\Omega_n(\xi)=\omega_n(\xi)/\omega_M$, where $\omega_M=\sqrt{K_M/M}$ is the natural frequency of the uncoupled master structure. The corresponding expression for a discrete set of oscillators becomes

$$\frac{\mu}{N}\sum_0^N\left|\frac{\Omega_n^2(\xi)}{1-\Omega_n^2(\xi)+j\Omega_n(\xi)\chi}\right|^2 d\xi-1=0. \tag{64}$$

As expected, the solution of Eq. (64) differs from that of the continuous case, Eq. (63).

B. Approximate solutions for χ

The solution of Eq. (63) requires the frequency distribution of the oscillators $\Omega_n(\xi)$ in the layer. In general, the characteristic equation (63) for χ does not admit closed-form solutions. However, integrating Eq. (63) near its peak at $\Omega_n(\xi_M)=1$ yields an approximate solution for particular frequency distributions.

Considering a distribution $\Omega_n(\xi)=\Omega_{\max}\xi^p$ as an example, where $\Omega_{\max}=(\omega_{\max}/\omega_M)$ represents the ratio of the highest frequency in the layer to the forcing frequency, changes the integration variable ξ to $\Omega_n(\xi)$ in Eq. (63):

$$\frac{\mu}{p\Omega_{\max}^{1/p}}\int_0^{\Omega_{\max}}\frac{1}{\Omega_n}\left|\frac{\Omega_n^2}{1-\Omega_n^2+j\Omega_n\chi}\right|^2 d\Omega_n-1=0. \tag{65}$$

An approximate evaluation of this integral around its peak as a product of an effective bandwidth $\chi\pi/2$ (viz., Ref. 19), and the peak value of the integrand $1/\chi^2$, produces

$$\chi=\frac{\pi}{2}\frac{\mu}{p\Omega_{\max}^{1/p}}. \tag{66}$$

This approximate result suggests that for a general distribution of oscillator frequencies, χ depends on μ and varies inversely with $\Omega_{\max}$. For the special case of linear frequency distribution $(p=1)$, the damping factor given by Eq. (66) becomes the same as published results.[7] For example, adapting the result for a damping coefficient $C_M=(\pi/2)\omega^2\,dm/d\omega$ given in, viz., Refs. 3, 7, 16, where $dm/d\omega$ is spectral distribution of mass along the layer, gives a form similar to the approximation in Eq. (66). Substituting for $dm/d\omega=m_L/\omega_{\max}$, where m_L is the total mass of the layer, produces

$$\chi=\frac{C_M}{\omega_M M}=\frac{\pi}{2\mu}\frac{\omega^2}{\omega_M\omega_{\max}}. \tag{67}$$

When evaluated at the natural frequency of the primary structure that forces the layer, Eq. (67) yields

$$\chi=\frac{\pi}{2}\frac{\mu}{\Omega_{\max}}. \tag{68}$$

Equations (68) and (66) give the same result for the linear case, $p=1$.

Figure 5 displays results obtained by numerically solving Eq. (63) for several different frequency distributions while keeping $\Omega_{\max}=2$. Results also confirm that χ varies in direct proportion to the mass ratio μ and inversely with $\Omega_{\max}$.

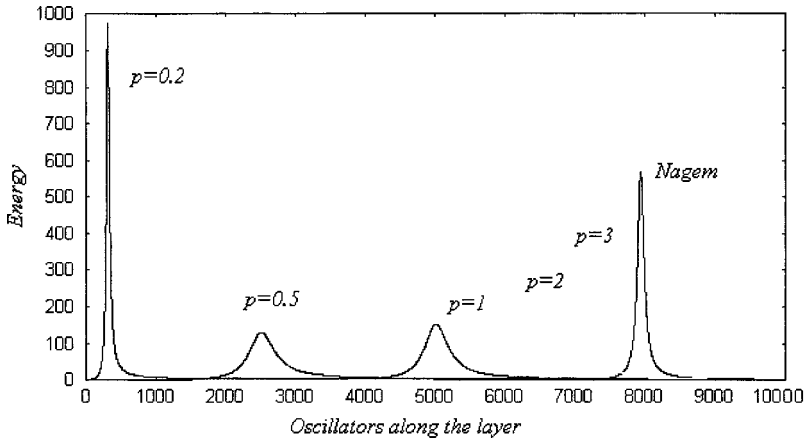


FIG. 6. The distribution of energy in the layer of 10 000 oscillators changes according to the distribution of natural frequencies of the oscillators.

V. ENERGY AND PHASE DISTRIBUTION IN THE LAYER

In this section we examine the energy and the phase distributions within the layer of oscillators along ξ to investigate how the energy absorbed by the layer distributes itself among the oscillators within it and the possible role of energy partitioning among the oscillators on apparent damping and return time.

A. Energy distribution

Substituting for $A(\xi)$ from Eq. (60) into the energy distribution along the layer:

$$E(\xi)=\frac{1}{2}k(\xi)A^2(\xi)=\frac{1}{2}m[\omega_n(\xi)A(\xi)]^2, \tag{69}$$

and normalizing the result as $e(\xi)=E(\xi)/(1/2)mV_0^2$ gives an expression for the nondimensional energy distribution along the layer:

$$e(\xi)=\frac{\Omega_n^4(\xi)}{|1-\Omega_n^2(\xi)+j\chi\Omega_n(\xi)|^2}. \tag{70}$$

An examination of energy density distribution expression Eq. (70) along the layer of oscillators reveals a ''spatial'' resonance between the natural frequency of the primary structure and the layer of oscillators. Energy distribution also shows that the energy concentrates mainly in the part of the layer tuned with ω_M, i.e., in the region $\Omega_n(\xi)\sim 1$. Energy distribution among the oscillators relates to the apparent damping in that increased values of damping correlate with energy acquisition by a wider part of the layer. Thus, the layer becomes a more effective energy sink when it is able to spread the trapped energy among more of its oscillators, suggesting that its effectiveness relates to the ability of energy partitioning. Also, because χ is proportional to the mass ratio μ, then μ must also affect energy partitioning. A numerical analysis of how $e(\xi)$ varies with $\omega_n(\xi)$ is shown in Fig. 6 for several cases of frequency distribution expressed by $\omega_n(\xi)=\omega_{\max}\xi^p$. For all the cases illustrated, $\omega_{\max}=2$. The results show that energy remains primarily around the oscillators approximately tuned with the master frequency.

B. Phase distribution

The phase distribution along the layer follows from the corresponding displacement expression:

$$\hat{x}(\xi,t)=\omega_n(\xi)\int_{-\infty}^{+\infty}x_M(t)e^{-j\omega_n(\xi)t}\,dt, \tag{71}$$

which with its modulus, $|\hat{x}(\xi,t)|$, given in Eq. (58), also represents the envelope of the oscillations as a function of distribution parameter ξ along the layer. The phase distribution at each point ξ within the layer can be found as

$$\angle\hat{x}(\xi,t)=\arctan\frac{\mathbf{Im}\{\hat{x}(\xi,t)\}}{\mathbf{Re}\{\hat{x}(\xi,t)\}}=\arctan\frac{\omega_n(\xi)C_M}{M\omega_n^2(\xi)-K_M}. \tag{72}$$

In non-dimensional form, the phase at any point ξ along the layer becomes

$$\angle\hat{x}(\xi,t)=\arctan\frac{\chi\Omega_n(\xi)}{[\Omega_n^2(\xi)-1]}. \tag{73}$$

This expression shows that oscillators have coherent phases only far outside of the resonance band. The widening of the resonance band and the decrease of phase coherence within it correspond to an increased value of χ. Thus, an optimum distribution of a layer of oscillators would have a broad resonance band within which coherence is minimal.

VI. NUMERICAL RESULTS

In this section we present numerical results to compare and validate the approximate expressions developed above for return times and apparent damping factor with the solutions obtained by direct simulations of the same.

Figures 7–8 present results in terms of nondimensional response $(x_M\omega_M/V_0)$ of the master structure as a function of nondimensional time $(t\omega_M)$ for different master frequencies and layer frequency distributions $\omega_n(\xi)=\omega_{\max}\xi^p$ by varying the value of p. In the simulations shown, the dynamic system consists of a master structure and $N=100$ parallel oscillators with a total mass ratio $\mu=m_L/M$. For comparison, the position of the predicted value of return times are indicated with a vertical line. They also include a plot of the approxi-

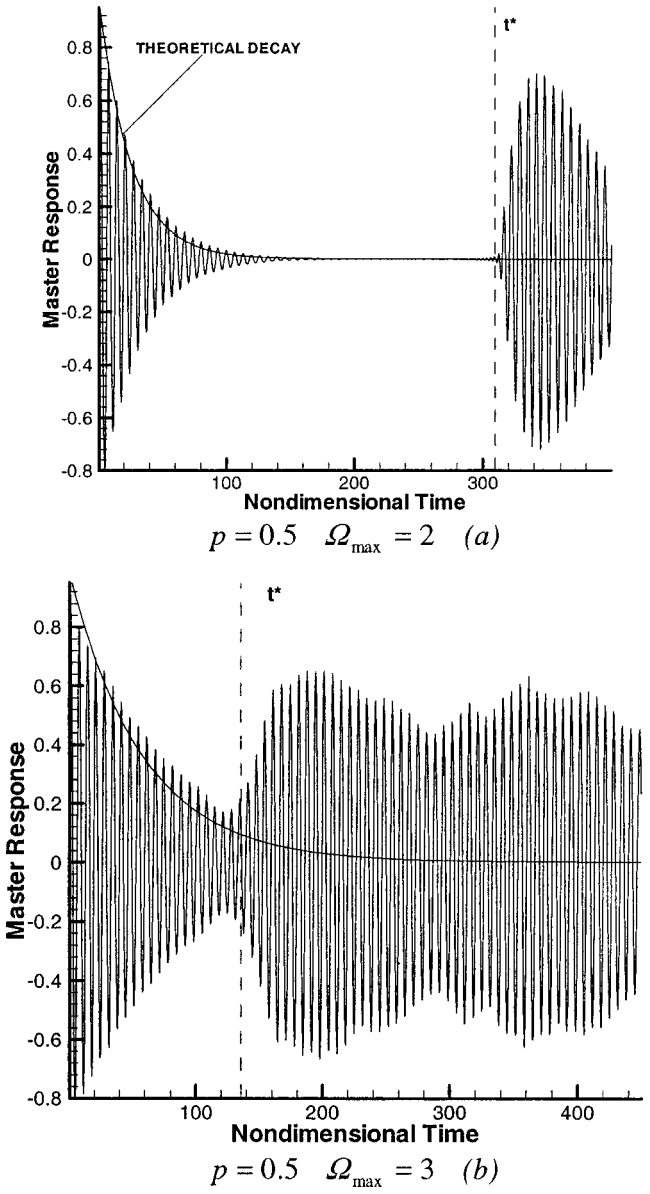


FIG. 7. Response of the primary structure to an impulse. $N=100$.

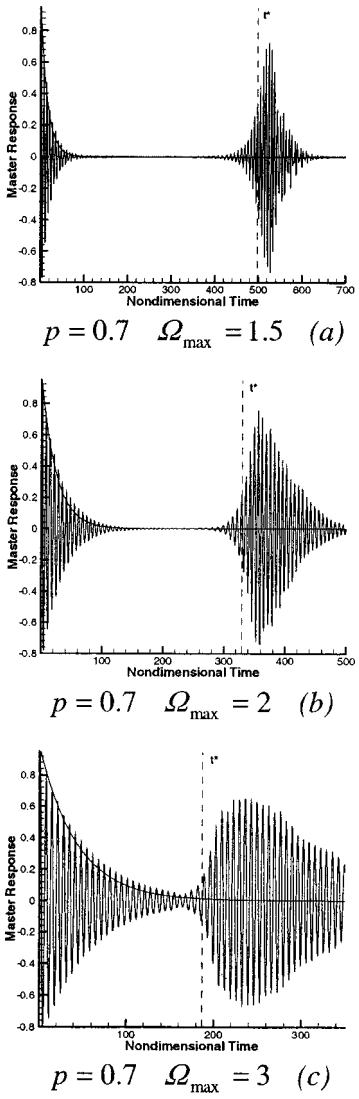


FIG. 8. Response of the primary structure to an impulse. $N=100$.

mate decay obtained with χ provided by Eq. (63) (or equivalently with the simpler expression provided in the next section).

Figures 7 and 8 show results for values $p<1$ and Figs. 9–11 show results for cases where $p>1$. Each case considers different values of the ratio of master frequency to the maximum value of the layer frequency, $\Omega_{\max}=\omega_{n\,\max}/\omega_M$. The predicted return time in Figs. 7–8 refers to the master wave or the primary wave group obtained with Eq. (49) and shows good agreement with the return time in the simulations based on with the numerical solution of Eq. (1) and Eq. (2). The results also show a very good agreement for the decay rate during the transient response of the primary structure as do the following cases.

The cases for $p>1$ presented in Figs. 9–11, both the pulse return time and the master return time are also indicated. The arrival of the pulse wave group is a small-amplitude disturbance that affects the regular harmonic response induced by the master, and it is slightly ahead of the master wave group.

Although the frequency distribution $\omega_n(\xi)=\omega_{\max}\xi^p$ used in the developments is sufficiently general, the wave analogy presented here applies to any arbitrary distribution of frequency along the layer, e.g., exponential.

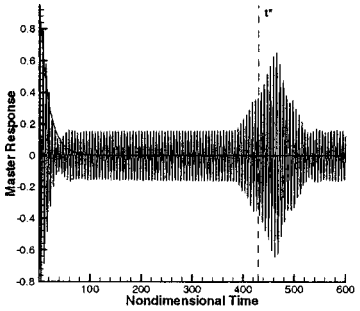
VII. RETURN TIME T^* AND APPARENT DAMPING χ

A relationship between return time t^* and the apparent damping factor χ results from χ in (66) and the nondimensionalized expression (49) for t_M^*:

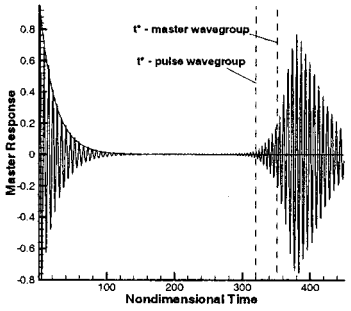
$$\frac{t_M^*}{T_M}=N\left(\frac{1}{p}\frac{1}{\Omega_{\max}^{1/p}}\right), \tag{74}$$

$$\chi=\frac{\pi}{2}\mu\left(\frac{1}{p}\frac{1}{\Omega_{\max}^{1/p}}\right). \tag{75}$$

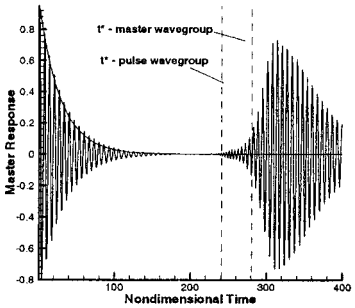
Combining these equations leads to a relationship between t_M^* and χ as



$p=1.3 \quad \Omega_{\max}=1.1 \quad (a)$



$p=1.3 \quad \Omega_{\max}=1.5 \quad (b)$



$p=1.3 \quad \Omega_{\max}=2 \quad (c)$

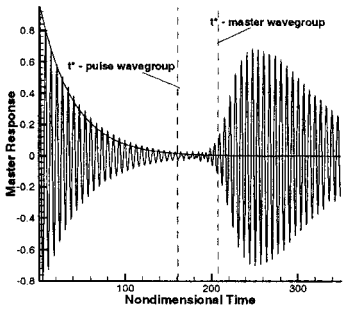
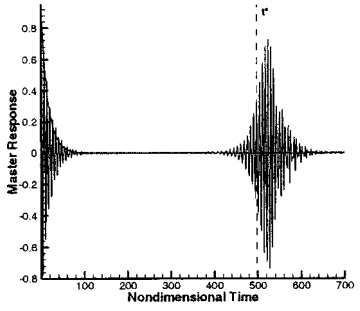


FIG. 9. Response of the primary structure to an impulse. $N=100$.

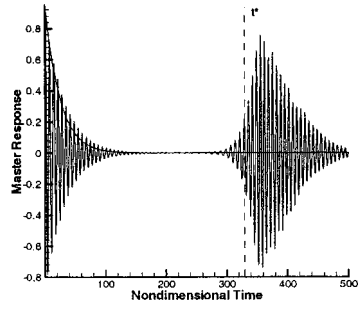
$$\chi=\frac{\pi}{2}\frac{\mu}{N}\frac{t_M^*}{T_M}. \tag{76}$$

Figure 12 shows a comparison of the apparent damping factor χ predicted by Eq. (76) and that obtained by solving Eq. (63), indicating a good agreement between them. It appears from Fig. 12 that a sharp maximum develops in χ (and also in t_M^*) when it is examined as a function of p. By using the last three equations, the maximum values of both χ and t_M^* are obtained for

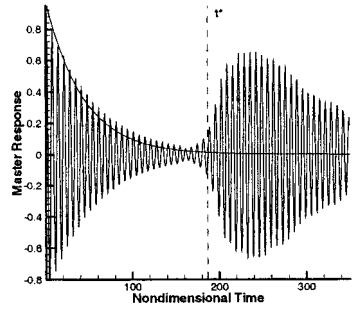
$$p=\ln \Omega_{\max}.$$



$p=0.7 \quad \Omega_{\max}=1.5 \quad (a)$



$p=0.7 \quad \Omega_{\max}=2 \quad (b)$



$p=0.7 \quad \Omega_{\max}=3 \quad (c)$

FIG. 10. Response of the primary structure to an impulse. $N=100$.

This result provides a helpful element in view of a design rule for enhancing absorption by an attached set of oscillators.

VIII. CONCLUDING REMARKS

The analysis presented in this paper started with the reaction force by a set of oscillators attached to a primary structure, comparing the result with the corresponding case where the attached oscillators form a continuous set. The closed-form solution for the discrete set produced the return time, t^*. In the limiting case, the discrete solution yielded the force from a continuous set that showed a decrease of force amplitude with time.

Expressing the force from a continuous set as a sum of a finite discrete set and its remainder, a general expression was developed for the return time from a discrete set that has an arbitrary distribution of natural frequencies. This approach shows the significance of the frequency distributions and gives a lower bound for the return time.

Inspired by the similarity between the periodic return of energy from the oscillator set to the primary and reflections of waves in a waveguide, a wave analogy was developed.

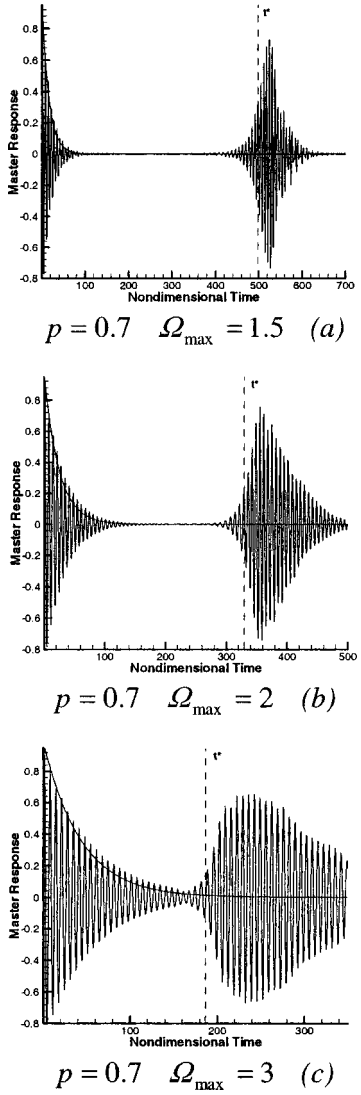


FIG. 11. Response of the primary structure to an impulse. $N = 100$.

This approach considers the relative phase among the oscillators in the set as traveling wave, and yields two types of group waves: one associated with the free vibrations of the oscillators only and the other with their forced vibrations at the natural frequency of the master. Results, given for a general frequency distribution, produce $t^* = 2\pi/\Delta\omega$ as the return time for the case of a linear frequency distribution.

Finally, the analysis concludes with a derivation of a simple expression for the apparent damping factor that indicates the initial energy transfer from the master to the oscillator following an impulsive excitation applied to the primary structure. The derivation uses the Hilbert envelope to describe the energy transfer between the primary and the attached set. The simple expression derived agrees extremely well with the transient part of the direct numerical simulations. The analysis concludes with a simple relationship between the damping factor and return time that yields an optimum value for a large damping factor or long return time in terms of the natural frequency distribution of the oscillators in the attached set.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. 0084581.

APPENDIX: DERIVATION OF EXPRESSION FOR REMAINDER

The integral of a function $f(\xi)$ that has continuous derivatives up to second order within an interval [0,1] can be expressed as

$$\int_0^1 f(\xi)d\xi = \sum_{i=1}^{N} f(\xi_i)\Delta\xi + \mathcal{R}(\Delta\xi), \tag{A1}$$

where, as before, $\Delta\xi = 1/N$, $\xi_i = i\,\Delta\xi$ and $\mathcal{R}(\Delta\xi)$ represents the remainder.

The remaining term $\mathcal{R}_i(\Delta\xi)$ associated with a single subinterval $[\xi_i, \xi_i + \Delta\xi]$ and its first derivative can be expressed as

$$\mathcal{R}_i(\Delta\xi) = \int_{\xi_i}^{\xi_i+\Delta\xi} f(\xi)d\xi - f(\xi_i)\Delta\xi, \tag{A2}$$

$$\mathcal{R}_i'(\Delta\xi) = f(\xi_i + \Delta\xi) - f(\xi_i). \tag{A3}$$

Applying the mean value theorem to Eq. (A3) gives

$$\mathcal{R}_i'(\Delta\xi) = f'(\overline{\xi_i})\Delta\xi, \quad \overline{\xi} \in [\xi_i, \xi_i + \Delta\xi]. \tag{A4}$$

Considering the identity

$$\mathcal{R}_i(\Delta\xi) = \mathcal{R}_i(0) + \int_0^{\Delta\xi} \mathcal{R}_i'(\xi)d\xi, \tag{A5}$$

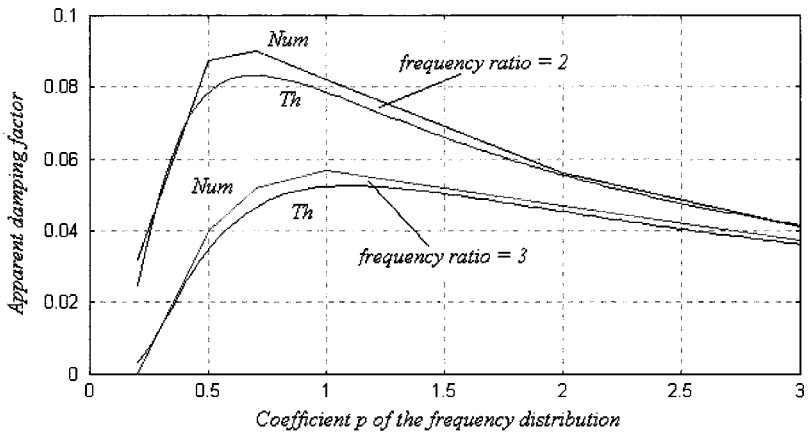


FIG. 12. The variation of the apparent damping factor with p.

and since from Eq. (A2), the first term, $\mathcal{R}_i(0)=0$, the substitution of Eq. (A4) to (A5) produces an expression for the remaining term in the subinterval $[\xi_i, \xi_i+\Delta\xi]$:

$$\mathcal{R}_i(\Delta\xi)=\frac{1}{2}\Delta\xi^2 f'(\bar{\xi}_i). \tag{A6}$$

A summation of expression (A6) over i produces the remainder for the entire interval [0,1],

$$\mathcal{R}(\Delta\xi)=\frac{1}{2}\Delta\xi^2\sum_{i=1}^{N} f'(\bar{\xi}_i). \tag{A7}$$

Since

$$f'_{\min}\leqslant\frac{1}{N}\sum_{i=1}^{N} f'(\bar{\xi}_i)\leqslant f'_{\max}, \tag{A8}$$

where $f'_{\min}$ and $f'_{\max}$ represent the minimum and maximum values of $f'(\xi)$ for $\xi\in[0,1]$. Since $f'(\xi)$ is a continuous function, it can take any value within $[f'_{\min}, f'_{\max}]$. Thus, there exists a particular value for which

$$f'(\bar{\xi})=\frac{1}{N}\sum_{i=1}^{N} f'(\bar{\xi}_i). \tag{A9}$$

The substitution in Eq. (A7) with $\Delta\xi=1/N$ results in the expression for the remainder used in Eq. (18),

$$\mathcal{R}(\Delta\xi)=\frac{1}{2}N\Delta\xi^2 f'(\bar{\xi})=\frac{1}{2}\Delta\xi f'(\bar{\xi}). \tag{A10}$$

[1] C. E. Celik and A. Akay, ''Dissipation in solids: Thermal oscillations of atoms,'' J. Acoust. Soc. Am. **108**, 184–191 (2000).

[2] C. Soize, ''A model and numerical method in the medium frequency range for vibroacoustic predictions using the theory of structural fuzzy,'' J. Acoust. Soc. Am. **94**, 849–865 (1993).

[3] A. D. Pierce, V. W. Sparrow, and D. A. Russel, ''Fundamental structural–Acoustic idealization for structure with fuzzy internals,'' J. Vibr. Acoust. **117**, 339–348 (1995).

[4] M. Strasberg and D. Feit, ''Vibration damping of large structures induced by attached small resonant structures,'' J. Acoust. Soc. Am. **99**, 335–344 (1996).

[5] M. Strasberg, ''Continuous structure as ''fuzzy'' substructures,'' J. Acoust. Soc. Am. **100**, 3456–3459 (1996).

[6] R. J. Nagem, I. Veljkovic, and G. Sandri, ''Vibration damping by a continuous distribution of undamped oscillators,'' J. Sound Vib. **207**, 429–434 (1997).

[7] R. L. Weaver, ''The effect of an undamped finite degree of freedom 'fuzzy' substructure: Numerical solutions and theoretical discussion,'' J. Acoust. Soc. Am. **100**, 3159–3164 (1996).

[8] R. L. Weaver, ''Mean and mean-square response of a prototypical master/fuzzy structure,'' J. Acoust. Soc. Am. **101**, 1441–1449 (1997).

[9] R. L. Weaver, ''Multiple-scattering theory for mean responses in a plate with sprung masses,'' J. Acoust. Soc. Am. **101**, 3466–3474 (1997).

[10] R. L. Weaver, ''Mean-square response in a plate with sprung masses, energy flow and diffusion,'' J. Acoust. Soc. Am. **103**, 414–427 (1998).

[11] R. L. Weaver, ''Equipartition and mean-square response in large undamped structures,'' J. Acoust. Soc. Am. **110**, 894–903 (2001).

[12] G. Maidanik and K. J. Becker, ''Noise control of a master harmonic oscillator coupled to a set of satellite harmonic oscillators,'' J. Acoust. Soc. Am. **104**, 2628–2637 (1998).

[13] G. Maidanik and K. J. Becker, ''Characterization of multiple-sprung masses for wideband noise control,'' J. Acoust. Soc. Am. **106**, 3109–3118 (1999).

[14] G. Maidanik and K. J. Becker, ''Criteria for designing multiple-sprung mass for wideband noise control,'' J. Acoust. Soc. Am. **106**, 3119–3127 (1999).

[15] G. Maidanik, ''Induced damping by a *nearly* continuous distribution of *nearly* undamped oscillators: Linear analysis,'' J. Sound Vib. **240**, 717–731 (2001).

[16] K. Xu and T. Igusa, ''Dynamic characteristics of multiple substructures with closely spaced frequencies,'' Earthquake Eng. Struct. Dyn. **21**, 1050–1070 (1992).

[17] M. V. Drexel and J. H. Ginsberg, ''Modal overlap and dissipation effects of a cantilever beam with multiple attached oscillators,'' J. Vibr. Acoust. **123**, 181–187 (2001).

[18] A. Carcaterra and A. Sestieri, ''Complex envelope displacement analysis: A quasi-static approach to vibrations,'' J. Sound Vib. **201**, 205–233 (1997).

[19] S. H. Crandall and W. D. Mark, *Random Vibration in Mechanical Systems* (Academic, New York, 1963).

Experimental validation of a nonparametric probabilistic model of nonhomogeneous uncertainties for dynamical systems

Hamid Chebli
Structural Dynamics and Coupled Systems Department, ONERA, 92322 Châtillon, France

Christian Soize[a)]
Laboratoire de Mécanique, Université de Marne-La-Vallée, 77454 Marne-La-Vallée, France

(Received 4 April 2003; revised 14 November 2003; accepted 17 November 2003)

The paper deals with an experimental validation of a nonparametric probabilistic model of nonhomogeneous uncertainties for dynamical systems. The theory used, recently introduced, allows model uncertainties and data uncertainties to be simultaneously taken into account. An experiment devoted to this validation was specifically developed. The experimental model is constituted of two simple dural rectangular plates connected together with a complex joint. In the mean mechanical model, the complex joint, which is constituted of two additional plates attached with 40 screw-bolts, is modeled by a homogeneous orthotropic continuous plate with constant thickness, as usual. Consequently, the mean model introduces a region (the joint) which has a high level of uncertainties. The objective of the paper is to present the experiment and the comparisons of the theoretical prediction with the experiments. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1639335]

PACS numbers: 43.40.Hb, 43.60.Cg, 43.55.Cs [RLW] Pages: 697–705

I. INTRODUCTION

In order to improve the robustness of the prediction models of complex dynamical systems, random uncertainties have to be taken into account. There are two main types of uncertainties: (1) *model uncertainties* which are induced by the modeling process which allows the mathematical model of the real dynamical systems to be constructed and (2) *data uncertainties* which correspond to the errors on the parameters of the constructed model. It is well known that probabilistic parametric approaches are very efficient to model the data uncertainties (for instance, see Refs. 1–7). However, probabilistic approaches do not allow model uncertainties to be taken into account. This is the reason why a nonparametric probabilistic model of random uncertainties for dynamical systems has recently been proposed.[8,9] This nonparametric approach allows model uncertainties and data uncertainties to be simultaneously taken into account. Without re-explaining all the details of this nonparametric approach, hereafter, we summarize the principal idea. First, it should be noted that any probability measure can be constructed using the maximum entropy principle[9] with appropriate constraints. Consequently, when a probability measure has to be constructed, the problem is to utilize the constraints defined by the better available information. For dynamical systems, the most important available information is relative to the operators of the nominal model (such as the mass, damping, and stiffness operators of the nominal model presently called the mean model) and to their algebraic properties. Modeling random uncertainties leads the corresponding operators of the dynamical system to be random and to verify the same algebraic properties (positiveness, invertibility, etc.). For the actual dynamical system, the corresponding operators are unknown but it can be stated that these operators have to verify the same algebraic properties which are absolutely general for any dynamical system. In addition, for the statistical identification of the predictive model, the actual dynamical system under consideration has to be considered as a realization of a random dynamical system. For a given operator of the predictive model, the probabilistic parametric approach of data uncertainties yields an operator range which is a subset $\mathbb{S}_{\text{par}} \subset \mathbb{S}$ of an appropriate functional space $\mathbb{S}$ and which can be too ''small'' due to the presence of additional model uncertainties. This means that, if the model uncertainties are large, the corresponding operator of the actual dynamical system does not belong to subset $\mathbb{S}_{\text{par}}$. The proposed probabilistic nonparametric approach consists in constructing a ''bigger subset'' $\mathbb{S}_{\text{nonpar}}$ using only the known general algebraic properties which hold for any dynamical system. Consequently, by construction, $\mathbb{S}_{\text{nonpar}}$ contains subset $\mathbb{S}_{\text{par}}$ (that is to say $\mathbb{S}_{\text{par}} \subset \mathbb{S}_{\text{nonpar}}$) and then, the operator of the actual dynamical system belongs to $\mathbb{S}_{\text{nonpar}}$. Then, a probability measure is constructed on $\mathbb{S}_{\text{nonpar}}$ and this is done using the maximum entropy principle with the constraints defined by all the available information related to the given operator of the dynamical system.

In Refs. 8 and 9, the theory is presented for master dynamical systems with homogeneous random uncertainties. Generally, for complex dynamical systems, random uncertainties are not homogeneous. The level of uncertainty is different from a part to another one. For instance, if we consider two simple structures connected together with a complex joint, the uncertainties of the mean mechanical model are important in the part constituted of the joint and are small in the two parts constituted of the two simple structures. Consequently, the nonparametric probabilistic model of random uncertainties has been extended to the case of nonhomogeneous uncertainties by using a substructuring

[a)]Electronic mail: soize@univ-mlv.fr

technique.[10–12] Such an approach combines the Craig–Bampton dynamic substructuring method[13] and the nonparametric probabilistic model.

In order to validate the nonparametric probabilistic model of nonhomogeneous uncertainties for complex dynamical systems, an experiment has specifically been carried out.[14] In this paper, we present this experiment and the experimental validation of the theory developed in Refs. 8–12 and which constitutes the first experimental validation of the proposed nonparametric probabilistic approach.

The experimental setup is made up of two dural rectangular plates connected together by a complex joint constituted of two additional plates and 40 screw bolts. Then, the dynamical system has three natural subdomains. The first one is a simple rectangular plate, the second one is a complex joint, and the third one is an another simple rectangular plate. In order to evaluate the role played by the value and by the distribution of the screwing-couples, 21 configurations corresponding to 21 values and distributions of the screwing couples have been tested.

The mean numerical model is constituted of a finite element model of a simplified schematization of the dynamical system for which the complex joint is replaced by an equivalent simple orthotropic continuous plate with constant thickness. The modeling uncertainties are induced by the introduction of such a simple schematization of the complex joint. The finite element model of the orthotropic continuous plate modeling the complex joint has been updated using the first seventh experimental elastic modes. Nevertheless, in spite of this dynamical updating, the mean finite element model cannot correctly predict the frequency response functions on a broad frequency band, especially when the frequency increases. The comparisons of the experimental frequency response functions with the mean finite element model clearly show that the mean model is good for the low-frequency range but is not able to predict the experimental responses for higher frequencies. The introduction of a nonparametric model of nonhomogeneous random uncertainties allows the robustness of the prediction to be improved.

In Sec. II, the experimental setup is presented. Section III deals with the mean mechanical model and its finite element modeling. In Sec. IV, we present the comparisons of the frequency responses given by the mean model with the experimental frequency responses. Section V is devoted to the comparisons of the random frequency responses of the stochastic model with the experimental frequency responses. In the following, FRF means ''frequency response function.''

II. EXPERIMENT

The experiment which is briefly described in the following was performed at ONERA.[14] The system studied is constituted of two simple rectangular thin plates connected by a bolted joint, as explained in Sec. I.

A. Description of the experimental setup

The two simple plates of the experimental system are denoted P_1 and P_3. Plate P_1 is a rectangular dural plate with thickness 0.003 m, width 0.40 m, and length 0.60 m. Plate P_3 is a rectangular dural plate with thickness 0.003 m, width 0.50 m, and length 0.60 m. Plates P_1 and P_3 are connected on their length by a joint constituted of two rectangular plates with thickness 0.002 m, width 0.14 m, and length 0.60 m (see Fig. 1) and are attached with two lines of 20 screw bolts. The experimental dynamical system is hanged with a very low eigenfrequency. The excitation is done with an electrodynamic shaker located in plate P_3 (see k_{exit} in Fig. 1). The response is identified with 29 accelerometers. In the following, we present experimental measurements for the two accelerometers located at point k_{exit} (driving force) in plate P_3 and at point k_{obs} in plate P_1 (see Fig. 1).

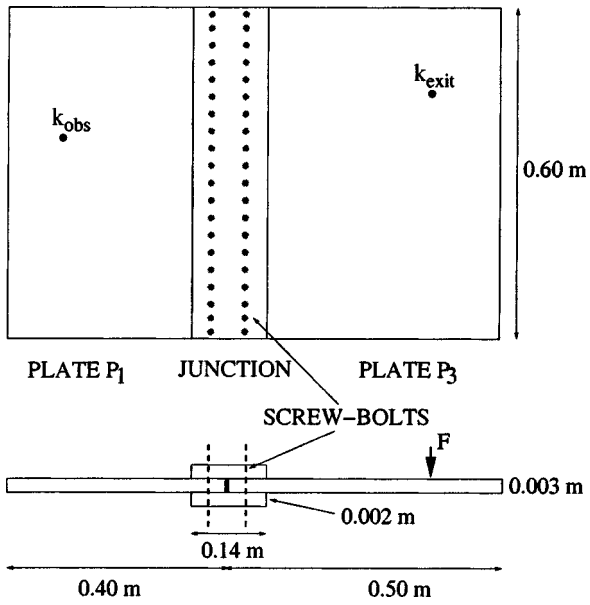


FIG. 1. Experimental mechanical system.

B. Experimental configurations

In order to study the influence of the screw-bolt prestresses, 21 experimental configurations were defined. Each experimental configuration corresponds to a given value of each screw-bolt prestress for the 40 screw bolts. The prestresses of these 40 screw bolts correspond to the realization of a random screwing couple of each screw bolt. The probability distribution is uniform on the interval $[2,5]$ N×m. The 21 experimental configurations were tested.

C. Experimental measurements

For each experimental configuration $j=1,...,21$, the data collected are the accelerations $\{A_k^{\exp,j}(f),k=1,...,29\}$, for f in $B=[20,2000]$ Hz. From the measured accelerations, the displacements $U_k^{\exp,j}(f)$ are easily deduced. For each f in B, we introduce the functions $f\mapsto\max_k^{\exp}(f)$ and $f\mapsto\min_k^{\exp}(f)$ related to the 21 experimental configurations and defined by

$$\max_k^{\exp}(f)=\max_{j=1,...,21}\{10\log_{10}(|U_k^{\exp,j}(f)|^2)\}, \tag{1}$$

$$\min_k^{\exp}(f)=\min_{j=1,...,21}\{10\log_{10}(|U_k^{\exp,j}(f)|^2)\}. \tag{2}$$

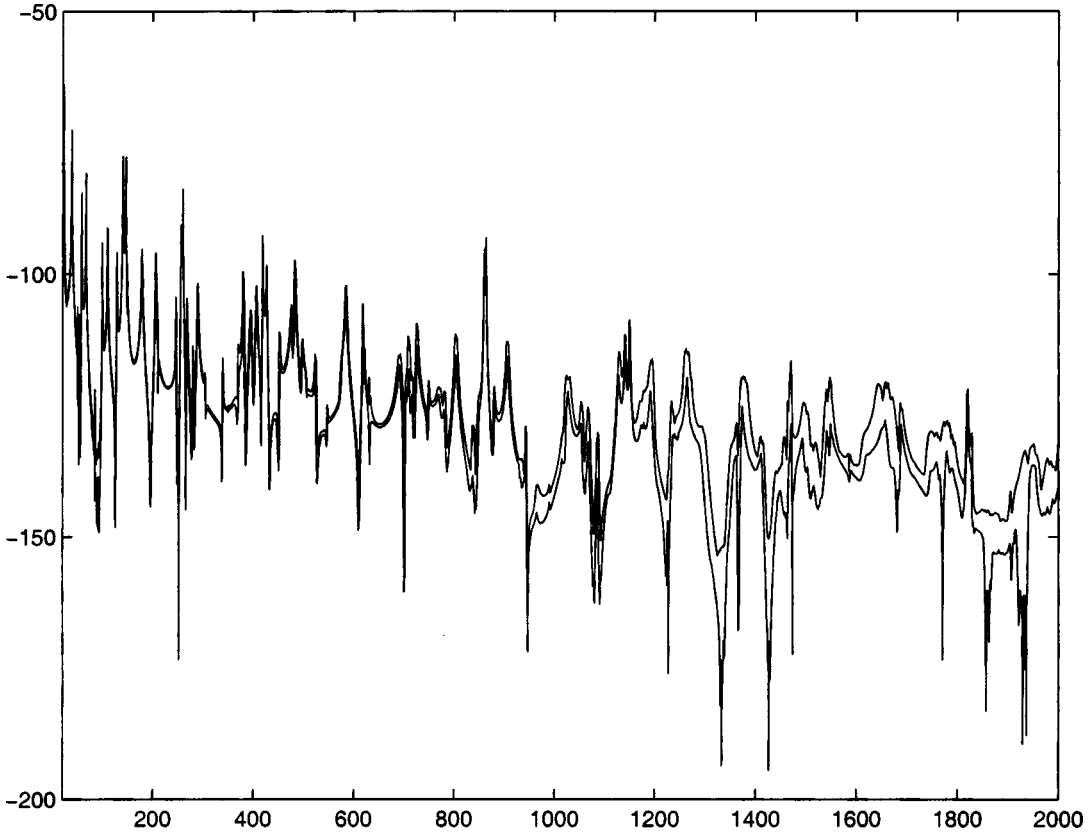


FIG. 2. Experimental frequency response function at observation point k_{obs}: graphs of functions $f \mapsto \max_{k_{\text{obs}}}^{\text{exp}}(f)$ (upper solid line) and $f \mapsto \min_{k_{\text{obs}}}^{\text{exp}}(f)$ (lower solid line).

D. Influence of the screw-bolt prestresses

Figure 2 displays the experimental results of the displacement FRF modulus for observation point k_{obs} in plate P_1. The upper and lower curves correspond to the graphs of functions $f \mapsto \max_{k_{\text{obs}}}^{\text{exp}}(f)$ and $f \mapsto \min_{k_{\text{obs}}}^{\text{exp}}(f)$, respectively. It can be seen that the dispersion induced by the random screw-bolt prestresses is completely negligible in the [0,200] Hz frequency band, is small in [200,800] Hz, and is still limited in the [800,2000] Hz frequency band. It can be concluded that the screw-bolt prestresses induce a small dispersion which is increasing with the frequency.

III. MEAN MECHANICAL MODEL OF THE EXPERIMENTAL CONFIGURATIONS

This section deals with the construction of the mean mechanical model of the experimental configurations and with the updating of the finite element model.

A. Mean mechanical model

We consider the complete structure Ω whose dimension is 0.9 m$\times$0.6 m in a Cartesian coordinate system $(OXYZ)$ defined in Fig. 3. This structure is constituted of three substructures Ω^1 (plate P_1), Ω^2 (joint), and Ω^3 (plate P_3), whose lengths are 0.33, 0.14, and 0.43 m, respectively. The thickness of substructures Ω^1 and Ω^3 is 0.003 m. The joint is modeled by substructure Ω^2 whose thickness is 0.007 m (corresponding to 0.002+0.003+0.002, see Fig. 3). The coupling interface between Ω^1 and Ω^2 is Σ^1, and between Ω^2 and Ω^3 is Σ^2.

Plates P_1 (substructure Ω^1) and P_3 (substructure Ω^3) are made in dural and consequently are modeled by a homogeneous isotropic material whose measured characteristics are the following: mass density 2800 kg/m^3, Young's modulus 7.05×10^{10} N/m^2, Poisson's rate 0.33, and damping rate

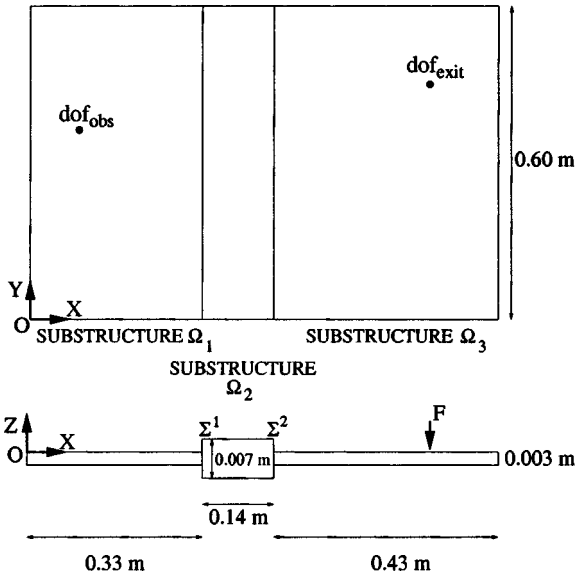


FIG. 3. Mean mechanical model of the experimental mechanical system.

0.2%. If the joint (substructure Ω^2) is modeled by a homogeneous and isotropic plate with thickness 0.007 m, then the seven lowest eigenvalues of the structure Ω can never been predicted with a sufficient accuracy when the prediction is compared with the experiments. It can be concluded that an orthotropic medium is required for the mean model of the joint. Such an orthotropic model was developed and the characteristic properties were identified by updating the finite element model (see Sec. III C). The mass of each accelerometer (0.0007 kg) is introduced in the model. An appropriate system of four springs models the experimental hanging.

The modal damping rates have been experimentally identified for the ten first elastic modes. The mean value of these experimental damping rates is 0.2%. For the mean damping of each substructure, a hysteretic model is used. Consequently, the mean damping matrix of each substructure is proportional to the mean stiffness matrix with a coefficient equal to $2\beta/(2\pi f)$ in which $\beta=0.002$ and where f is the frequency in Hz.

The coordinates of excitation point k_{exit} are (0.77,0.45,0). In the frequency domain, the driving force is a constant normal force applied to the plate P_3 (substructure Ω^3) over frequency band B and is equal to 1. The normal displacement at points k_{exit} and k_{obs} (corresponding to the location of two accelerometers) is observed to characterize the dynamical responses of structure Ω. The coordinates of observation point k_{obs} are (0.09,0.36,0).

B. Mean finite element model of the experimental configurations

The mean finite element models of the three substructures Ω^1, Ω^2, and Ω^3 are constructed with compatible meshes constituted of four-nodes bending plate elements. The mesh size is 0.01 m$\times$0.01 m. The assembled structure Ω has 16 653 DOFs, the number of internal DOFs of Ω^1 (plate P_1), Ω^2 (joint), and Ω^3 (plate P_3) is 6039, 2379, and 7869, respectively. The number of coupling DOFs for each interface Σ^1 and Σ^2 is 183. The observation DOFs corresponding to the normal displacement at driving point k_{exit} and at observation point k_{obs} are denoted DOF_{exit} and DOF_{obs}, respectively (see Fig. 3).

C. Properties of the mean model

The mesh size of the mean finite element model has been adapted to the wave numbers of the highest mode in the frequency band $B=[0,2000]$ Hz of analysis. The modal density of the mean structure is such that there are 8 modes in the frequency band [20,100] Hz, 27 modes in [100,500] Hz, 29 modes in [500,1000] Hz, 33 modes in [1000,1500] Hz and finally, 29 modes in [1500,2000] Hz. For the mode corresponding to the highest eigenfrequency in frequency band B (the highest mode), the corresponding wave numbers n and m in X and Y directions (see Fig. 3) are $n=15$ and $m=10$. Since mean finite element mesh is such that there are 90 finite element nodes in X and direction and 60 finite element nodes in Y direction, the spatial sampling for the highest mode is 6 nodes per ''half wavelength'' in each direction. Consequently, the mean finite element mesh is sufficiently refined to give converged results in the sense of the mean model.

D. Updating the mean finite element model

In order to improve the quality of the mean finite element model, the mean finite element model of the joint (substructure Ω^2) has been updated with respect to the seven lowest eigenfrequencies which have been measured. The updated model for substructure Ω^2 corresponds to an orthotropic material whose Young's moduli are $E_X=1.90\times10^{10}\,\text{N/m}^2$ and $E_Y=7.40\times10^{10}\,\text{N/m}^2$ and whose in-plane shear modulus is $1.23\times10^{10}\,\text{N/m}^2$. Moreover, the mass of each screw bolt is taken into account and the resulting mass density of substructure Ω^2 is equal to 3086 kg/m^3. The seven lowest experimental eigenfrequencies are 20.35, 22.12, 37.97, 52.80, 57.99, 66.58, and 84.70 Hz, while the calculated eigenfrequencies with the updated mean finite element model are 20.35, 22.54, 40.25, 53.32, 57.10, 66.36, and 84.02 Hz.

IV. RESPONSE OF THE MEAN MODEL AND COMPARISONS WITH THE EXPERIMENTAL RESULTS

In this section, the response of the mean model calculated by using the Craig–Bampton dynamic substructuring method is compared with the experimental results. Figures 4 and 5 are related to the driving point k_{exit} and to the observation point k_{obs}, respectively. Each figure is relative to the FRF modulus in dB for the displacement and compares the mean numerical predictions with the experimental upper and lower envelopes. Although the updating of the mean finite element model using the experiments has been carried out, for frequencies greater than 400 Hz, Figs. 4 and 5 clearly show that the updated mean model is not sufficiently accurate to predict the experimental response for frequencies greater than 400 Hz. Figures 7 and 9 can be used as a zoom (magnified figures) for studying frequencies less than 400 Hz and show that, in fact, the updated mean finite element model is not sufficiently accurate to predict the experimental response for frequencies greater than 100 Hz. It should be noted that the number of DOFs for such a system is relatively high with respect to the usual state of art for this type of finite element modeling. Since the effect of the screwing couple is under control through the experimental configurations, it is clear that the main source of uncertainty does not come from this phenomena, but from the model errors (model uncertainties) whose effects increase with frequency (errors induced by the schematization used for modeling a complex joint by a simple continuous plate). Data errors (data uncertainties) exist as well.

V. NONPARAMETRIC MODEL OF NONHOMOGENEOUS RANDOM UNCERTAINTIES AND COMPARISONS WITH THE EXPERIMENTAL RESULTS

As explained in Sec. IV, model uncertainties associated with the mean model constructed in Sec. III have to be taken

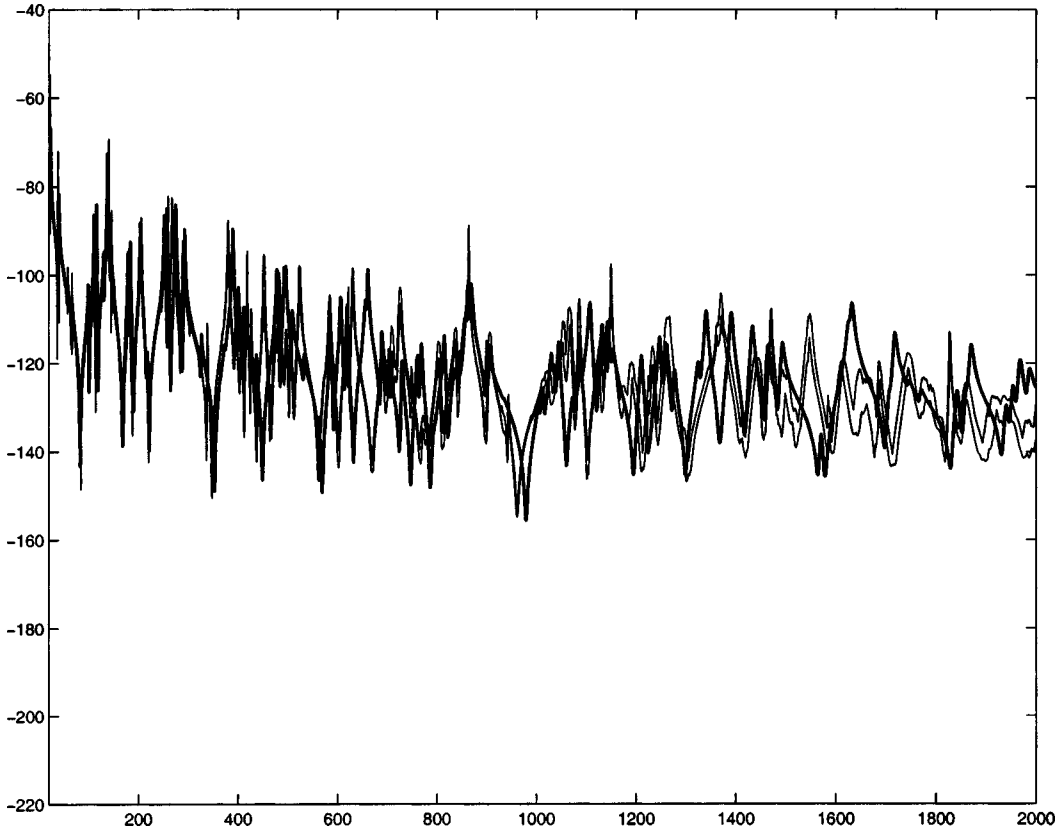


FIG. 4. Experimental comparisons at driving point k_{exit} over frequency band [20,2000] Hz: numerical prediction of the frequency response function modulus in dB (thick solid line) and, experimental graphs of $f \mapsto \max_{k_{\text{exit}}}^{\text{exp}}(f)$ (upper thin solid line) and $f \mapsto \min_{k_{\text{exit}}}^{\text{exp}}(f)$ (lower thin solid line).

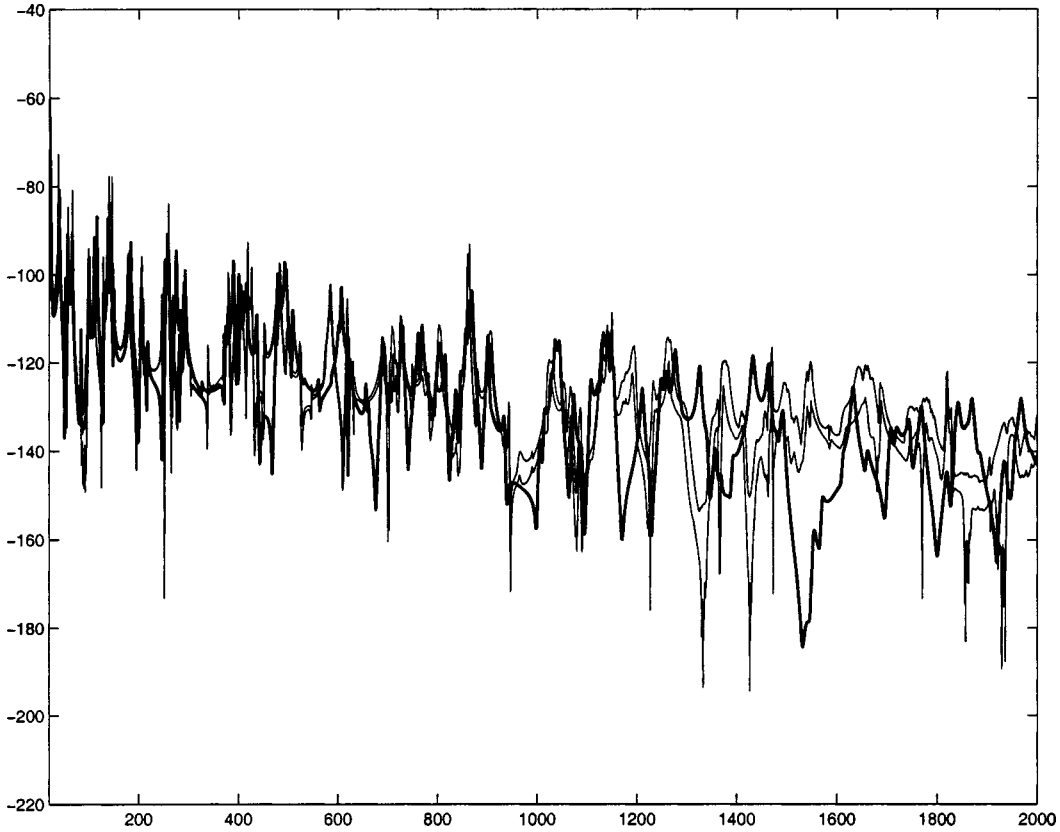


FIG. 5. Experimental comparisons at observation point k_{obs} over frequency band [20,2000] Hz: numerical prediction of the frequency response function modulus in dB (thick solid line) and, experimental graphs of $f \mapsto \max_{k_{\text{obs}}}^{\text{exp}}(f)$ (upper thin solid line) and $f \mapsto \min_{k_{\text{obs}}}^{\text{exp}}(f)$ (lower thin solid line).

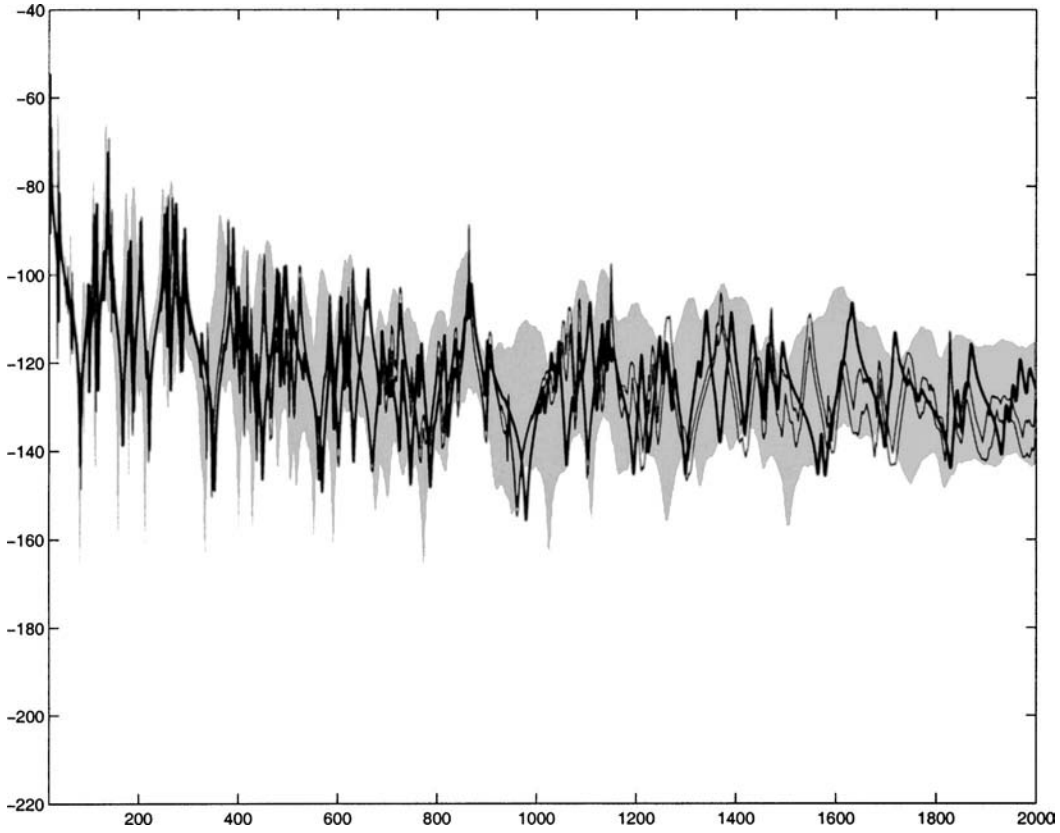


FIG. 6. Experimental comparisons at driving point k_{exit} over frequency band [20,2000] Hz: confidence region prediction in dB with the probabilistic numerical model (gray region), numerical prediction in dB of the mean model (thick solid line), and experimental graphs of $f \mapsto \max_{k_{\text{exit}}}^{\text{exp}}(f)$ (upper thin solid line) and $f \mapsto \min_{k_{\text{exit}}}^{\text{exp}}(f)$ (lower thin solid line).

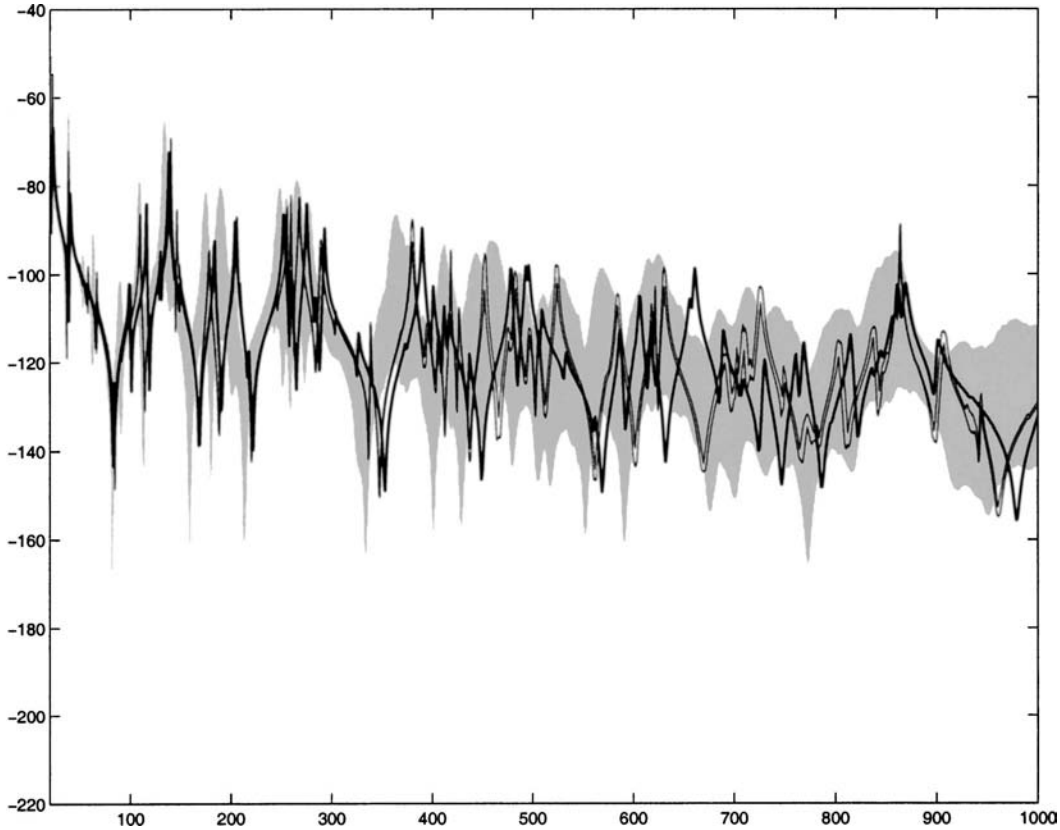


FIG. 7. Experimental comparisons at driving point k_{exit} over frequency band [20,1000] Hz: confidence region prediction in dB with the probabilistic numerical model (gray region), numerical prediction in dB of the mean model (thick solid line) and experimental graphs of $f \mapsto \max_{k_{\text{exit}}}^{\text{exp}}(f)$ (upper thin solid line) and $f \mapsto \min_{k_{\text{exit}}}^{\text{exp}}(f)$ (lower thin solid line).

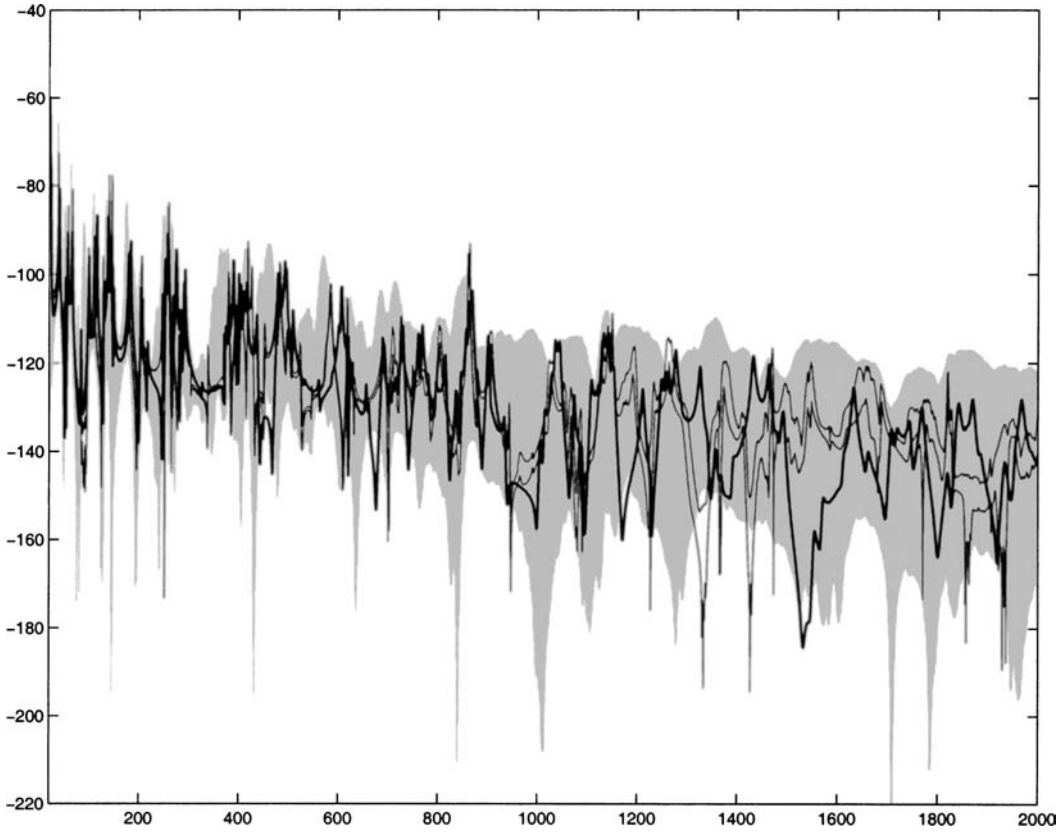


FIG. 8. Experimental comparisons at observation point k_{obs} over frequency band [20,2000] Hz: confidence region prediction in dB with the probabilistic numerical model (gray region), numerical prediction in dB of the mean model (thick solid line) and experimental graphs of $f \mapsto \max_{k_{\text{obs}}}^{\text{exp}}(f)$ (upper thin solid line) and $f \mapsto \min_{k_{\text{obs}}}^{\text{exp}}(f)$ (lower thin solid line).

into account to increase the robustness of the prediction due to the unsatisfaying prediction given by the mean model. The dispersion parameter δ which allows the nonparametric probabilistic model to be controled is defined in Ref. 9 and is a dimensionless parameter belonging to interval [0,1]. Such a parameter is similar to the variation index in statistics ($\delta=0$ corresponds to no uncertainties and $\delta=1$ to the biggest possible level of uncertainty). Moreover, the level of uncertainty is more important in the substructure modeling the joint. In order to model such nonhomogeneous random uncertainties, the nonparametric probabilistic method proposed in Refs. 10–12 is used.

Although the goal of this paper is not to present some methods for identifying δ parameters, hereinafter, we summarize the principle possible methods: (i) Since such a nonparametric model of random uncertainties allows model errors to be taken into account, δ parameters can be used as global parameters to carry out a sensitivity analysis with respect to the model errors and data errors. With such an approach, δ is not fixed but has to belong to an interval. This method has been used in this paper for calibrating the values of δ parameters for the two plates and for the joint. (ii) A second method consists in performing an experimental identification using an appropriate inverse method. Such research is in progess. (iii) A third method consists in calibrating δ parameters by using a parametric approach. An example of such a method is given in Ref. 15. (iv) A fourth method consists in constructing an estimation of δ for a given type of substructures (for instance for a given type of joints) using any one of the two above-mentioned methods.

A. Definition of the uncertainty level

A sensitivity analysis has been carried out with respect to the dispersion parameters δ_M, δ_D, and δ_K related to the mass, damping, and stiffness uncertainties in each substructure. Concerning the details of this analysis, we refer the reader to Ref. 10. From this analysis, it can be deduced that a low level of uncertainties has to be introduced in substructures Ω_1 and Ω_3, and a high level of uncertainties has to be used for substructure Ω_2 modeling the joint. Since the mass distribution of the mean model is relatively well known (with respect to frequency band B), we have taken $\delta_{M^1}=\delta_{M^2}=\delta_{M^3}=0$ for the three substructures. Finally, the optimal values for the δ_D and δ_K parameters (corresponding to a confidence region of the stochastic response associated with a probability of 0.95) are $\delta_{D^1}=\delta_{D^3}=0.10$, $\delta_{D^2}=0.80$, and $\delta_{K^1}=\delta_{K^3}=0.15$, $\delta_{K^2}=0.80$. These values correspond to a high level of uncertainties in substructure Ω^2 (the joint) for the stiffness and the damping. Since the mass of the mean model is updated with the experiment, no significant errors exist for the mass. It should be noted that if a moderate level of uncertainty is considered for substructure Ω^2, then it is not sufficient to obtain a confidence region in which the experimental results are. In addition, it seems to be inconsistent to increase the level of uncertainty in substructures Ω^1 and

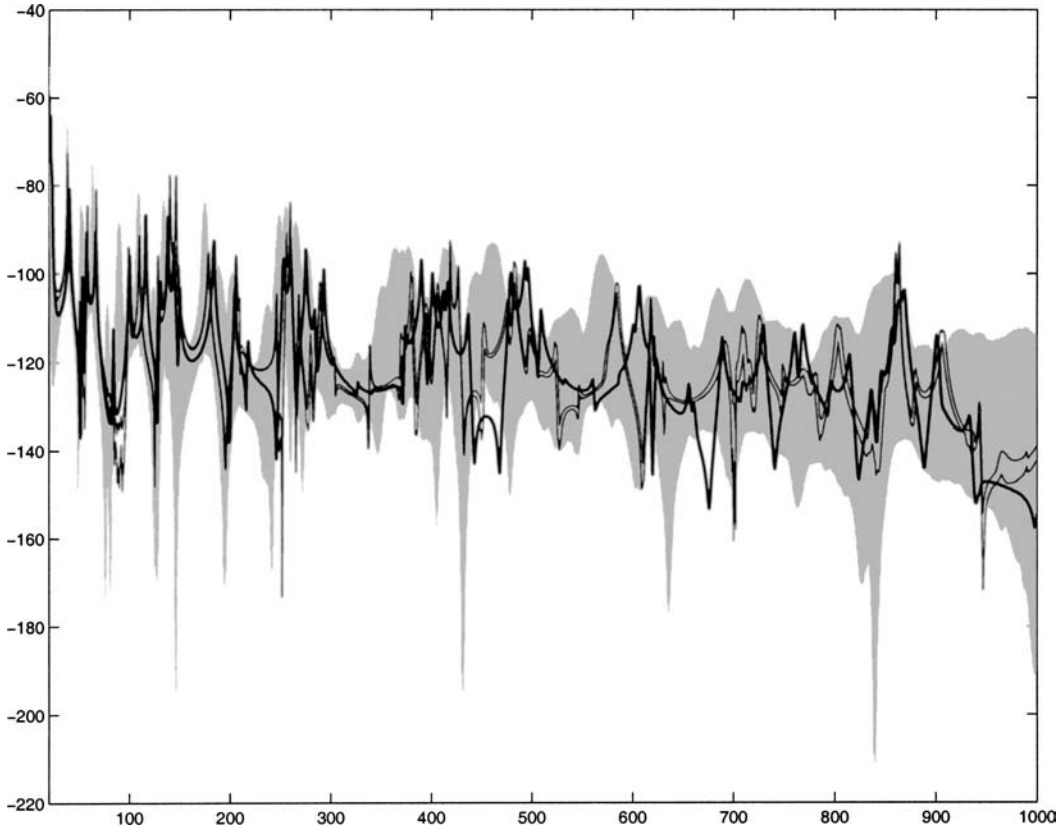


FIG. 9. Experimental comparisons at observation point k_{obs} over frequency band [20,1000] Hz: confidence region prediction in dB with the probabilistic numerical model (gray region), numerical prediction in dB of the mean model (thick solid line) and experimental graphs of $f \mapsto \max_{k_{\text{obs}}}^{\text{exp}}(f)$ (upper thin solid line) and $f \mapsto \min_{k_{\text{obs}}}^{\text{exp}}(f)$ (lower thin solid line).

Ω^3 modeling plates P_1 and P_3, but substructure Ω^2 has to be considered with a high level of uncertainty.

B. Comparisons: calculation-measurement

Let $\mathbf{U}(f)=(U_1(f),\ldots,U_n(f))$ be the random response of structure Ω. A confidence region, constructed by using the Tchebychev inequality, is defined by the upper and lower envelopes $f \mapsto dB_k^+(f)$ and $f \mapsto dB_k^-(f)$ of the FRF modulus in dB for the displacement and corresponding to a given probability level P_c such that

$$P\{dB_k^-(f) < dB_k(f) < dB_k^+(f)\} \geqslant P_c, \tag{3}$$

in which the random variable $dB_k(f)$ is defined by $dB_k(f) = 10\log_{10}(|U_k(f)|^2)$ and where the lower envelope $dB_k^-(f)$ and the upper envelope $dB_k^+(f)$ are such that

$$dB_k^-(f) = 2dB_k^0(f) - dB_k^+(f), \tag{4}$$

$$dB_k^0(f) = 20\log_{10}(|E\{U_k(f)\}|), \tag{5}$$

$$dB_k^+(f) = 20\log_{10}\{|E\{U_k(f)\}| + a_k(f)\},$$

$$a_k(f) = \frac{\sigma_{U_k(f)}}{\sqrt{1-P_c}}, \tag{6}$$

in which $E\{U_k(f)\}$ and $\sigma_{U_k(f)}$ are the mean value and the standard deviation of the complex-valued random variable $U_k(f)$. Figures 6–9 display the comparisons between the experiment and the prediction. It should be noted that the mean value $E\{U_k(f)\}$ and the standard deviation $\sigma_{U_k(f)}$ can easily be deduced from Eqs. (4) to (6). Figures 6 and 7 are relative to the displacement at driving point k_{exit} over frequency band [20,2000] Hz (see Fig. 6) and [20,1000] Hz (zoom given by Fig. 7). Figures 8 and 9 are relative to the displacement at observation point k_{obs} over frequency band [20,2000] Hz (see Fig. 8) and [20,1000] Hz (zoom given by Fig. 9). On each figure, it can be seen (1) the gray region whose upper and lower envelopes are defined by the graphs $f \mapsto dB_k^+(f)$ and $f \mapsto dB_k^-(f)$ and which corresponds to the confidence region with $P_c = 0.95$, constructed using the probabilistic numerical model, (2) the thick solid line corresponding to the response of the mean finite element model and (3) the upper and lower thin solid lines corresponding to the upper and lower envelopes of the experimental FRF for the displacements. It can be seen that the gray region contains almost all the experimental results (in the 0.95 probability) and that the bandwith of this region increases with the frequency. This is consistent with the fact that the role played by uncertainties increase with the frequency.

In order to explain the efficiency of the proposed nonparametric probabilistic method of nonhomogeneous model uncertainties, we focus the analysis given in the following on a particular aspect of this experimental comparison. Let us consider the resonance predicted by the mean finite element

model at frequency 660 Hz (see the peak of the thick solid line in Fig. 7). The corresponding experimental resonance is given at frequency 628 Hz (see the peak of the upper and the lower thin solid lines in Fig. 7). If the mean model was perfectly correct, these two resonances should be located at the same frequency. It is not the case. It can be seen that the gray region is not centered around the response of the mean finite element model but is shifted allowing the confidence region to include the experimental resonance at 628 Hz. Such an analysis can be reproduced for almost all the differences appearing between the experimental results and the prediction over all the frequency band.

VI. CONCLUSIONS

We have presented an experimental validation of the nonparametric model of random uncertainties proposed in Ref. 9 for the case of nonhomogeneous uncertainties induced by a complex joint. The experimental comparisons show the efficiency of the theory proposed and show that such a theory has to predict a confidence region which has to increase with the frequency, that is the case for this theory. Finally, this paper shows that, in a broad low-frequency range, the use of such a nonparametric probabilistic model allows the dynamical responses of a structure including a complex joint to be predicted by using a usual mean mechanical model of the joint.

ACKNOWLEDGMENT

The authors thank ONERA which has supported this research.

[1] E. J. Haug, K. K. Choi, and V. Komkov, *Design Sensitivity Analysis of Structural Systems* (Academic, New York, 1986).
[2] R. A. Ibrahim, ''Structural dynamics with parameter uncertainties,'' Appl. Mech. Rev. **40**, 309–328 (1987).
[3] E. Vanmarcke and M. Grigoriu, ''Stochastic finite element analysis of simple beams,'' J. Eng. Mech. **109**, 1203–1214 (1983).
[4] M. Shinozuka and G. Deodatis, ''Response variability of stochastic finite element systems,'' J. Eng. Mech. **114**, 499–519 (1988).
[5] P. D. Spanos and R. G. Ghanem, ''Stochastic finite element expansion for random media,'' J. Eng. Mech. **115**, 1035–1053 (1989).
[6] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach* (Springer, New York, 1991).
[7] M. Kleiber, D. H. Tran, and T. D. Hien, *The Stochastic Finite Element Method* (Wiley, New York, 1992).
[8] C. Soize, ''A nonparametric model of random uncertainties for reduced matrix models in structural dynamics,'' Probab. Eng. Mech. **15**, 277–294 (2000).
[9] C. Soize, ''Maximum entropy approach for modeling random uncertainties in transient elastodynamics,'' J. Acoust. Soc. Am. **109**, 1979–1996 (2001).
[10] H. Chebli, ''Modélisation des incertitudes aléatoires non homogénes en dynamique des structures pour le domaine des basses fréquencies,'' ''Nonhomogenous random uncertainties modeling in structural dynamics for the low-frequency domain,'' Doctoral Thesis from Conservatoire National des Arts et Métiers, 2002.
[11] C. Soize and H. Chebli, ''Random uncertainties model in dynamic substructuring using a nonparametric probabilistic model,'' Probab. Eng. Mech. **129**, 449–457 (2003).
[12] H. Chebli and C. Soize, ''Analyse vibratoire par sous-structuration avee modéle non paramétrique d'incertitudes aléatoires non homogénes,'' ''Vibration analysis by substructing technique with a nonparametric model of nonhomogeneous random uncertainties,'' Revue Européenne des Eléments Finis **11**, 234–246 (2002).
[13] R. R. Craig, Jr. and M. C. C. Bampton, ''Coupling of substructures of dynamic analyses,'' AIAA J. **14**, 1313–1319 (1968).
[14] M. Menelle and P. Baroin, ONERA Technical Report RT 1/05809 DDSS (2001).
[15] C. Soize, ''Random matrix theory and non-parametric model of random uncertainties,'' J. Sound Vib. **263**, 893–916 (2003).

Response variance prediction in the statistical energy analysis of built-up systems

Robin S. Langley[a)]
Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom

Vincent Cotoni
Vibro-Acoustic Sciences Inc., 12555 High Bluff Drive, Suite 310, San Diego, California 92130

(Received 18 April 2003; accepted for publication 24 November 2003)

In the statistical energy analysis (SEA) of high frequency noise and vibration, a complex engineering structure is represented as an assembly of subsystems. The response of the system to external excitation is expressed in terms of the vibrational energy of each subsystem, and these energies are found by employing the principle of power balance. Strictly the computed energy is an average taken over an ensemble of random structures, and for many years there has been interest in extending the SEA prediction to the variance of the energy. A variance prediction method for a general built-up structure is presented here. Closed form expressions for the variance are obtained in terms of the standard SEA parameters and an additional set of parameters α_k that describe the nature of the power input to each subsystem k, and α_{ks} that describe the nature of the coupling between subsystems k and s. The theory is validated by comparison with Monte Carlo simulations of plate networks and structural-acoustic systems. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1642621]

PACS numbers: 43.40.Qi, 43.20.Ks, 43.40.Dx [RLW] Pages: 706–718

I. INTRODUCTION

Statistical energy analysis (SEA) has been used for many years to predict the response of complex engineering systems to high frequency excitation (see, for example, Lyon, 1975; Lyon and DeJong, 1995; Craik, 1996). The main advantage of the method over more conventional displacement based approaches, such as the finite element method, is that relatively few degrees of freedom are involved: typically the system is modeled in terms of the energy levels of tens of subsystems rather than the displacement of tens or hundreds of thousands of nodal points. The method predicts the ''mean'' energy level in each of the subsystems, in the sense that the system is considered to have random properties—thus, for example, the method would predict the mean interior noise level for a fleet of cars manufactured on the same production line. There has long been interest in extending this prediction to the variance of the energy, so that a measure of the scatter in performance can be established, and this provides the motivation for the present work.

Much previous work has considered the statistics of the response of a single subsystem, such as a plate or acoustic volume, beginning with early work on room acoustics (for example, Schroeder, 1962), developing into Poisson process based work on structural components or acoustic volumes (for example, Lyon, 1969; Davy, 1986), and culminating in work employing statistical concepts based on the Gaussian orthogonal ensemble (GOE) (for example, Weaver, 1989; Lobkis *et al.* 2000; Langley and Brown, 2003a, b). The Poisson and GOE approaches involve different assumptions regarding the statistics of the natural frequencies of the random system: the former assumes that the natural frequencies constitute a Poisson point process, while the latter assumes that they conform to the statistics of the eigenvalues of a particular type of random matrix (Mehta, 1991). The Poisson assumption was originally introduced primarily for analytical convenience, and it was later discovered that the GOE assumption tends to be more physically sound. In fact, there is a large literature concerning the applicability or otherwise of the GOE model to dynamic systems, and in brief there is much numerical and experimental evidence to suggest that the model is appropriate in many cases—notable references are Weaver (1989), Ellegaard *et al.* (2001), Bertlesen *et al.* (2000), as well as numerous references quoted in the text by Mehta (1991). One requirement for the GOE to apply is that the system does not have ''symmetries''—indeed, if the system has many symmetries and the natural frequencies from all symmetry classes are superposed, then the Poisson model is found to apply. This condition is found in very special cases such as an exactly rectangular simply supported plate, or an exactly box-shaped room, but any perturbation of the system will disturb the symmetry and tend to produce GOE statistics. Another condition for the GOE to apply is that the system is sufficiently random for ''statistical overlap'' to occur, i.e., the random variation in an individual natural frequency exceeds the mean frequency spacing, so that veering between neighboring natural frequencies can occur. In view of this, the GOE would not normally be expected to apply to the lowest natural frequencies of a system, and Soize (2003) has recently investigated this issue (among others) for the first 30 modes of a particular example. Although deviations from the GOE were found, the distribution of the spacings between successive natural frequencies was shown to be well matched to the GOE up to a level of two or three times the

a)Electronic mail: rsl21@eng.cam.ac.uk

mean spacing. More generally, the degree of statistical overlap can be expected to increase with increasing frequency, so that the GOE is most suited to the study of high frequency dynamics involving the higher modes of vibration. In summary there is much evidence to suggest the applicability of the GOE to the natural frequencies of a random system, and response variance estimates based on the GOE have been found to show good agreement with simulations and experiments (Langley and Brown, 2003a, b) for a randomly perturbed *single* subsystem.

The most comprehensive analysis to date of response variance in a *built-up* system is that contained in the text by Lyon and DeJong (1995). In this approach the SEA equations are assumed to hold, but the coefficients in the equations and the power inputs to the various subsystems are taken to be random. A perturbation expansion then leads to an expression for the variance of the subsystem energies in terms of the variance of the equation coefficients and the power inputs. The latter variances are calculated via a heuristic extension to analytical results arising from Poisson point process theory. Experience has shown (Lyon, 2001) that the equations presented by Lyon and DeJong (1995), although containing much useful physical insight, can yield poor quantitative estimates of the response variance, and so an alternative approach is needed. In the present work, recent results regarding the response of a single subsystem (Langley and Brown, 2003a, b) are extended to the case of a built-up system. It is assumed that the uncoupled natural frequencies of each subsystem (sometimes termed the ''blocked'' natural frequencies in SEA) are random and conform to GOE statistics (Mehta, 1991). By using a perturbation approach similar to that of Lyon and DeJong (1995), but informed by exact energy-power relationships (Langley, 1989), a surprisingly straightforward closed form result for the energy variance of each subsystem is found. The theory is validated by comparison with Monte Carlo simulations for a range of example structures.

In what follows the SEA equations are summarized in Sec. II A, following which the exact energy-power relations that hold for a deterministic system are reviewed in Sec. II B. Exact and approximate energy-power relations for a random system are then considered in Sec. II C, leading to the development of an expression for the response variance in Sec. III. Numerical validation studies are reported in Sec. IV, followed by conclusions in Sec. V.

II. ENERGY-POWER EQUATIONS

A. SEA equations

Before considering variance issues, it is useful to review the relations that are postulated by statistical energy analysis (SEA) regarding the power input to a system and the vibrational energy stored in the various components. In SEA the system is represented as an assembly of subsystems, and the time-averaged energy of the jth subsystem for harmonic vibration of frequency ω is written as E_j. The subsystem is considered to be resonant, so that the total energy (kinetic plus potential) E_j is equal to twice the kinetic energy. The external power input to this subsystem is denoted by $P_{\mathrm{in},j}$, and a consideration of power balance leads to a set of equations of the form (Lyon and DeJong, 1995)

$$\omega \eta_j E_j + \sum_k \omega \eta_{jk} n_j (E_j/n_j - E_k/n_k) = P_{\mathrm{in},j}. \tag{1}$$

Here η_j is the subsystem loss factor, so that the first term on the left of Eq. (1) represents the power loss through internal dissipation. The second term represents the power loss to other subsystems, being based on the hypothesis that the power transfer between subsystems j and k is proportional to the difference in the modal energy levels: the term n_j is the modal density (i.e., the number of natural frequencies in a unit frequency band), so that E_j/n_j is a measure of the energy per mode. The terms η_{jk} are known as the coupling loss factors, and they can often be computed by considering energy transfer via elastic or acoustic waves at the junctions between subsystems (Lyon and DeJong, 1995).

Many authors have considered the conditions that must be met for Eq. (1) to apply (see, for example, Woodhouse, 1981; Keane and Price, 1987; Langley, 1989; Mace, 1994). If standard results for the coupling loss factors are employed, then one key point is that E_j should be interpreted as the *ensemble average* of the subsystem energy, taken over a collection of systems with random properties. Sometimes the response energy is averaged over a frequency band, and strictly in this case E_j should be interpreted as the ensemble average of the frequency averaged energy. It follows that in general the energy variable in Eq. (1) can be written in the form

$$E_j = \mathrm{E}\left[(1/\Delta) \int_{\omega-\Delta/2}^{\omega+\Delta/2} 2T_j \, \mathrm{d}\omega \right] = \mathrm{E}[2\overline{T}_j], \tag{2}$$

where E[ ] represents the ensemble average, Δ is the width of the frequency averaging band, T_j is the kinetic energy of the subsystem, and $\overline{T}_j$ is the band-averaged kinetic energy. It is sometimes argued that the energy is ergodic, in the sense that a frequency-band average taken on any particular deterministic realization of the system will yield the same result as an ensemble average; in this case E_j in Eq. (1) can be simply interpreted as a frequency average. The conditions for ergodicity have not been rigorously established, but they are likely to include the requirement that Δ is wide enough to encompass several resonant modes in each subsystem. In what follows Eq. (2) will be adopted, so that no assumptions are made regarding the ergodicity of the energy.

Analogous to Eq. (2), the input power that appears in Eq. (1) can be written as

$$P_{\mathrm{in},j} = \mathrm{E}\left[(1/\Delta) \int_{\omega-\Delta/2}^{\omega+\Delta/2} P_j \, \mathrm{d}\omega \right] = \mathrm{E}[\overline{P}_j], \tag{3}$$

where P_j is the power input prior to ensemble or frequency averaging, and $\overline{P}_j$ is the band-averaged power input.

The SEA equations represented by Eq. (1) can be written in the following matrix form,

$$\mathbf{C}\hat{\mathbf{E}} = \mathbf{P}_{\mathrm{in}}, \tag{4}$$

where $\hat{E}_j = E_j/n_j$ and

$$C_{jj}=\omega\left(\eta_j+\sum_{k\neq j}\eta_{jk}\right)n_j, \tag{5}$$

$$C_{jk}=-\omega\eta_{jk}n_j, \quad j\neq k. \tag{6}$$

For a given power input Eq. (4) can be solved to yield the mean energy, as defined by Eq. (2). The aim of the following analysis is to extend SEA by deriving an expression for the variance of the energy, and the starting point for this is to consider an exact relationship between the input power and the subsystem energies for a deterministic system.

B. Exact equations for a deterministic system

As in SEA, any deterministic dynamic system can be considered to be an assembly of subsystems. If each subsystem is subjected to external harmonic forcing of frequency ω, then it is possible to solve the equations of motion to yield both the power input to each subsystem and the vibrational energy of each subsystem. It is of interest to consider the conditions under which there exists an exact linear relationship between the power input and the energy in the form

$$\mathbf{AT}=\mathbf{P}, \tag{7}$$

where T_j is the kinetic energy of the jth subsystem, P_j is the power input to the subsystem, and $\mathbf{A}$ is an $N\times N$ square matrix, where N is the number of subsystems. Progress can be made by noting that if the detailed equations of motion of the whole system involve M degrees of freedom, then the external forces can be described by a vector of dimension $M\times 1$, $\mathbf{F}$ say. It can be noted that N and M will typically have very different magnitudes: a detailed finite element model of an automotive structure, for example, may have several million degrees of freedom, while a SEA model of the system may employ tens of subsystems, so that $N\ll M$. The dynamic displacement of the system will be harmonic with frequency ω and will depend linearly on the entries of the force vector $\mathbf{F}$ (in detail, the response vector is given by the inverse of the system dynamic stiffness matrix times the force vector). Now both $\mathbf{T}$ and $\mathbf{P}$ are second order in the system displacement, and hence each will have a time-average component and a zero mean component at frequency 2ω, arising from the fact that the product of any two first order components can be written in the form $\mathrm{Re}[a\exp(\mathrm{i}\omega t)]\times\mathrm{Re}[b\exp(\mathrm{i}\omega t)]=\mathrm{Re}[ab^*+ab\exp(2\mathrm{i}\omega t)]/2$. The concern here is with the time-average component, and it follows that since the displacement is a linear function of $\mathbf{F}$, then the time-constant components of both $\mathbf{T}$ and $\mathbf{P}$ must be second order in $\mathbf{F}$, and hence linear functions of the entries S_{rs} of the force cross-spectral matrix, defined as

$$\mathbf{S}=\mathbf{FF}^{*\mathrm{T}}. \tag{8}$$

The subsystem energies and power inputs can thus be expressed in the form

$$T_j=\sum_{r,s=1}^{M}a_{rs}S_{rs}, \tag{9}$$

$$P_j=\sum_{r,s=1}^{M}b_{rs}S_{rs}, \tag{10}$$

for some set of coefficients a_{rs} and b_{rs} determined by the system dynamic stiffness matrix. If the forces are random, rather than deterministic and harmonic, then Eq. (8) should be interpreted as a cross-spectral density matrix at the frequency ω, but otherwise the following analysis is unaltered. Now in order to produce Eq. (7) it must be possible to eliminate the terms S_{rs} from Eqs. (9) and (10) to leave a relation between $\mathbf{T}$ and $\mathbf{P}$ that involves only the system dependent terms a_{rs} and b_{rs}. However, there are M^2 components S_{rs} and Eqs. (9) and (10) each constitute only N independent equations. Taking into account the fact that $\mathbf{S}$ is a Hermitian matrix, there are actually $(M+1)M/2$ independent components S_{rs}, and hence it is not possible[1] to derive Eq. (7) from Eqs. (9) and (10) by eliminating $\mathbf{S}$ unless $N=(M+1)M/2$. In general $N\ll M$, and thus Eq. (7) *cannot* be derived without imposing restrictions on the cross-spectral matrix of the loading. To put this another way, it is not possible to find a matrix $\mathbf{A}$ that is valid for all possible loading conditions. However, it *is* possible to compute $\mathbf{A}$ for any restricted set of loads that have

$$\mathbf{S}=\sum_{n=1}^{N}s_n\mathbf{S}_n, \tag{11}$$

where s_n are N variables that describe the intensity of the loading, and $\mathbf{S}_n$ are N matrices that describe the spatial distribution of the loading. In this case a matrix $\mathbf{A}$ can be found that is valid for any set of intensity coefficients s_n, and this matrix will be a function of the entries of the matrices $\mathbf{S}_n$ as well as the system dependent terms a_{rs} and b_{rs}. A specific example of Eq. (11) would be to assume that the loading on each subsystem has a fixed spatial form but an arbitrary intensity; in this case s_n and $\mathbf{S}_n$ can be taken to relate to the loading on subsystem n.

The above argument can readily be extended to the case of frequency band-averaged energies and powers. It can be shown that providing the N force cross-spectral parameters s_n are constant over the frequency band, an exact equation in the form of Eq. (7) can be derived, so that

$$\mathbf{B}\overline{\mathbf{T}}=\overline{\mathbf{P}}, \tag{12}$$

where an over-bar represents a band-averaged quantity. The matrix $\mathbf{B}$ will tend to the matrix $\mathbf{A}$ as the width of the averaging band tends to zero, but in general the matrices will differ for a finite averaging bandwidth.

An insight into the structure of the matrix $\mathbf{B}$ can be gained by considering previous work on energy-power relationships. The entries of the matrix $\mathbf{B}$ are sometimes referred to as energy influence coefficients, and methods of computing $\mathbf{B}$ by using the finite element method have been discussed by Guyader *et al.* (1982), Fredö (1997), and Mace and Shorter (2000). From a more analytical point of view, a result in the form of Eq. (12) has been derived by Langley (1989) for the special case of rain-on-the-roof excitation of

each subsystem, which is a special case of Eq. (11) and a type of loading often assumed in SEA. Closed form solutions for $\mathbf{B}$ and $\overline{\mathbf{P}}$ were found in the form

$$\mathbf{B}=\lceil q\rfloor\mathbf{M}^{-1}, \tag{13}$$

$$\overline{P}_j=s_j q_j, \tag{14}$$

where $\lceil\ \rfloor$ represents a diagonal matrix, s_j is the intensity of the loading on subsystem j, and

$$q_j=\mathrm{Re}\left\{\int_\Omega\int_{V_j} i\omega G_{jj}(\mathbf{x},\mathbf{x},\omega)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\omega\right\}, \tag{15}$$

$$M_{jk}=(1/2)\int_\Omega\int_{V_j}\int_{V_k}\omega^2\rho_j|G_{jk}(\mathbf{x},\mathbf{y},\omega)|^2\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}\,\mathrm{d}\omega. \tag{16}$$

Furthermore, the Green function $G_{jk}(\mathbf{x},\mathbf{y},\omega)$ represents the response at location $\mathbf{x}$ on subsystem j to a point load of unit amplitude situated at location $\mathbf{y}$ on subsystem k, V_j and V_k are the spatial domains of subsystems j and k, ρ_j is the density of subsystem j, and $\Omega=[\omega-\Delta/2,\omega+\Delta/2]$ is the frequency averaging band. The matrix $\mathbf{B}$ is not symmetric in general, and Langley (1989) has derived an alternative symmetric version of Eq. (12) in the form

$$(1/\pi)\lceil q\rfloor\mathbf{M}_{\mathrm{sym}}^{-1}\lceil q\rfloor\{\pi\lceil\rho^{-1}q^{-1}\rfloor\overline{\mathbf{T}}\}=\overline{\mathbf{P}}, \tag{17}$$

where

$$M_{\mathrm{sym},jk}=(1/2)\int_\Omega\int_{V_j}\int_{V_k}\omega^2|G_{jk}(\mathbf{x},\mathbf{y},\omega)|^2\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}\,\mathrm{d}\omega. \tag{18}$$

The quantity q_j given by Eq. (15) is proportional to the space averaged point mobility of subsystem j, and it was shown by Langley (1989) that under certain conditions (mainly weak coupling and high modal overlap) the energy term that appears in Eq. (17) becomes

$$\{\pi\lceil\rho^{-1}q^{-1}\rfloor\overline{\mathbf{T}}\}_j=2\overline{T}_j/n_j, \tag{19}$$

which is precisely the energy variable used in SEA, Eq. (4), aside from ensemble averaging. The relationship between the symmetric matrix that appears on the left of Eq. (17) and the SEA matrix $\mathbf{C}$ was explored by Langley (1990), although this level of detail is not required here.

Equations (13)–(18) relate to the special case of rain-on-the-roof forcing, which will *not* be assumed *a priori* here, although these equations will be called upon when considering the basic structure of $\mathbf{B}$. The main *general* result to be noted is as follows: providing the cross-spectrum of the loading can be expressed in terms of only N independent parameters, then (*i*) Eq. (12) exists, and (*ii*) the matrix $\mathbf{B}$ is not necessarily symmetric. Equation (12) can be reexpressed in the form

$$(\tfrac{1}{2})\mathbf{B}\lceil n\rfloor\{\lceil n^{-1}\rfloor 2\overline{\mathbf{T}}\}=\overline{\mathbf{P}}\quad\text{or}\quad\mathbf{D}\overline{\mathbf{E}}=\overline{\mathbf{P}}, \tag{20}$$

where $\overline{E}_j=2\overline{T}_j/n_j$, and the (generally nonsymmetric) matrix $\mathbf{D}$ is defined accordingly. Equation (20) forms the basis of the following analysis. In can be noted that the SEA energy and power variables that appear in Eq. (4) are related to those that appear in Eq. (20) as follows:

$$\hat{\mathbf{E}}=\mathrm{E}[\overline{\mathbf{E}}], \tag{21}$$

$$\mathbf{P}_{\mathrm{in}}=\mathrm{E}[\overline{\mathbf{P}}]. \tag{22}$$

Thus, for subsystem j: $\overline{P}_j$ is the band averaged power input, $\overline{E}_j$ is the band averaged energy per mode, $P_{\mathrm{in},j}$ is the band and ensemble averaged power input, and $\hat{E}_j$ is the band and ensemble averaged energy per mode. It is acknowledged that this notation is slightly awkward: a cleaner notation could have been adopted by using E_j to represent the nonaveraged energy per mode, followed by consistent modifications to represent band and ensemble averaging. However E_j is routinely used in SEA, Eq. (1), to represent the ensemble and frequency averaged subsystem energy, and the present notation avoids a deviation from that convention. The implications of Eq. (20) regarding the statistics of the response of a random system are considered in the following section.

C. Exact and approximate equations for a random system

If the system of interest is random rather than deterministic, then it is appropriate to consider an ensemble of systems comprised of all possible random systems. Equation (20) will apply to any particular member of this ensemble, but the matrix $\mathbf{D}$ and the power input $\overline{\mathbf{P}}$ will vary over the ensemble. These quantities can be written as

$$\overline{\mathbf{P}}=\mathbf{P}_{\mathrm{in}}+\mathbf{P}_{\mathrm{ran}}, \tag{23}$$

$$\mathbf{D}=\mathbf{D}_0+\mathbf{D}_{\mathrm{ran}}, \tag{24}$$

where $\mathbf{P}_{\mathrm{in}}$ [see Eq. (22)] and $\mathbf{D}_0$ are the ensemble mean values, and $\mathbf{P}_{\mathrm{ran}}$ and $\mathbf{D}_{\mathrm{ran}}$ are the random fluctuations around these values. Equation (20) then yields for any member of the ensemble

$$\overline{\mathbf{E}}=(\mathbf{D}_0+\mathbf{D}_{\mathrm{ran}})^{-1}(\mathbf{P}_{\mathrm{in}}+\mathbf{P}_{\mathrm{ran}}). \tag{25}$$

If the system randomness is sufficiently small for second order products of $\mathbf{P}_{\mathrm{ran}}$ and $\mathbf{D}_{\mathrm{ran}}$ to be neglected, then Eq. (25) leads to the approximate result

$$\overline{\mathbf{E}}=\mathbf{D}_0^{-1}\mathbf{P}_{\mathrm{in}}+\mathbf{D}_0^{-1}\mathbf{P}_{\mathrm{ran}}-\mathbf{D}_0^{-1}\mathbf{D}_{\mathrm{ran}}\mathbf{D}_0^{-1}\mathbf{P}_{\mathrm{in}}. \tag{26}$$

By taking the expected value of Eq. (26), and noting Eq. (21), it follows that

$$\hat{\mathbf{E}}=\mathbf{D}_0^{-1}\mathbf{P}_{\mathrm{in}}. \tag{27}$$

Now a comparison with Eq. (4) demonstrates that

$$\mathbf{D}_0=\mathrm{E}[\mathbf{D}]=\mathbf{C}, \tag{28}$$

i.e., the expected value of the matrix that appears in the exact energy-power relationship is equal to the SEA matrix $\mathbf{C}$. Equation (28) has significant implications regarding the structure of the random matrix $\mathbf{D}_{\mathrm{ran}}$, as can be seen by summing the rows of Eq. (20) to yield a total power balance equation in the form

$$P_{\text{tot}} = \sum_j \overline{P}_j = \sum_j \sum_k D_{jk} \overline{E}_k$$

$$= \sum_j \omega \eta_j \overline{E}_j + \sum_j \sum_k D_{\text{ran},jk} \overline{E}_k . \tag{29}$$

Here η_j is the loss factor that appears in the SEA matrix, rather than the ''true'' loss factor of subsystem j, which might vary randomly over the ensemble. In the following analysis it is assumed initially that the loss factors are deterministic, so that η_j actually represents the true loss factor and $\mathbf{D}_{\text{ran}}$ is independent of the system damping. The result derived for the response variance can then be considered to be a conditional variance, and Baysian methods (i.e., the total probability theorem) can be used to compute the full variance under random loss factors, as discussed later in Sec. III E. Given deterministic loss factors, the first term on the right of Eq. (29) is the total power dissipated through damping and the term on the left is the total power input to the system; the fact that these two quantities must be equal leads to the conclusion

$$\sum_j \sum_k D_{\text{ran},jk} \overline{E}_k = 0 \Rightarrow \sum_j D_{\text{ran},jk} = 0, \tag{30}$$

where it has been noted that the first result must hold for all possible $\overline{E}_k$, hence implying the second result. Thus, according to Eq. (30), each column of $\mathbf{D}_{\text{ran}}$ must sum to zero, so that the entries of the matrix cannot have full statistical independence. This condition can be imposed by setting

$$D_{\text{ran},kk} = -\sum_{j \neq k} D_{\text{ran},jk} . \tag{31}$$

Equations (26) and (31) are used in what follows to derive an expression for the ensemble variance of the subsystem energies $\overline{E}_k$.

III. THE ENERGY VARIANCE

A. The variance equations

Equation (26) can be rewritten in the form

$$\overline{\mathbf{E}} - \hat{\mathbf{E}} = \mathbf{D}_0^{-1} \mathbf{P}_{\text{ran}} - \mathbf{D}_0^{-1} \mathbf{D}_{\text{ran}} \hat{\mathbf{E}}, \tag{32}$$

where it can be recalled that $\hat{\mathbf{E}} = \mathrm{E}[\overline{\mathbf{E}}]$ is the ensemble averaged energy predicted by SEA. If the entries of $\mathbf{P}_{\text{ran}}$ and $\mathbf{D}_{\text{ran}}$ are taken to be uncorrelated, apart from the constraint imposed by Eq. (31), then Eq. (32) yields

$$\mathrm{Var}(\overline{E}_j) = \sum_k (D_{0,jk}^{-1})^2 \mathrm{Var}(P_{\text{ran},k}) + \sum_k \sum_{s \neq k} [(D_{0,jk}^{-1} - D_{0,js}^{-1}) \hat{E}_s]^2 \mathrm{Var}(D_{\text{ran},ks}), \tag{33}$$

where Var( ) represents the variance, and $D_{0,jk}^{-1}$ represents the jk entry of $\mathbf{D}_0^{-1}$ (which is the inverse of the SEA matrix $\mathbf{C}$). The assumption regarding the lack of correlation of the entries of $\mathbf{P}_{\text{ran}}$ and $\mathbf{D}_{\text{ran}}$ is justified in what follows, and expressions are derived below for the variances that appear on the right-hand side of Eq. (33). The following analysis makes use of recent results for the statistics of the response of a single subsystem (Langley and Brown, 2003a, b), which are summarized in the Appendix. These results are based on the assumption that the system randomness is sufficiently strong to cause a high level of coupling between the modes of a nominally ''perfect'' structure [a condition sometimes referred to as high statistical overlap (Manohar and Keane, 1994)]. Under this condition the dynamic response statistics are independent of the details of the random variation in the physical properties, and depend only on gross parameters such as the subsystem modal overlap factors.

Equation (33) can be compared with the variance prediction derived by Lyon and DeJong (1995), which forms Eq. (12.3.3) of that reference. The key differences are that in Eq. (12.3.3) the matrix $\mathbf{D}$ (in the present notation) is assumed to be symmetric, and Eq. (30) is not enforced. Furthermore, the results employed for the variance terms on the right-hand side of the equation are based on Poisson natural frequency statistics in conjunction with a heuristic modification for the effect of band averaging. This contrasts with the approach below in which rigorous recent results based on the Gaussian orthogonal ensemble are employed. Thus the similarity in form between Eq. (12.3.3) and Eq. (33) rests solely on the fact that both equations are based on the first order expansion represented by Eq. (26). Having said this, the present analysis is much informed by the physical insights provided by Lyon and DeJong (1995) and the earlier work of Lyon (1969).

B. The variance of the input power

The term $\overline{P}_k$ represents the band-averaged input power to subsystem k, with mean value $P_{\text{in},k}$ and random part $P_{\text{ran},k}$ [see Eq. (23)]. The statistics of $\overline{P}_k$ can be derived by invoking the SEA assumption of ''weak coupling'' between the subsystems, which implies that the power input to subsystem k is dominated by the local dynamics of subsystem k. It can be noted that the choice and definition of subsystems within a SEA model is usually determined by an attempt to ensure weak coupling between the subsystems; failure to meet this condition can lead to a poor prediction of the mean energy (see, for example, Mace, 1994). For weak coupling the mean power input, $P_{\text{in},k}$, is the same as that which would be obtained were subsystem k in isolation rather than connected to the rest of the system. The power input to an isolated subsystem has exactly the form of Eq. (A1) [see, for example, Eqs. (14) and (15) with the Green function G_{jj} expressed in terms of the modes of the uncoupled subsystem], and hence the analysis reported in the Appendix can be used to estimate the input power variance. This yields

$$\mathrm{Var}(P_{\text{ran},k}) = P_{\text{in},k}^2 r^2(\alpha_k, m_k', B_k'), \tag{34}$$

where the function r^2 is given by Eqs. (A2) and (A6). The term α_k depends upon the nature of the loading, as explained in the Appendix, and m_k' and B_k' are effective modal overlap and bandwidth parameters defined by Eqs. (A4) and (A5). The latter two terms depend on the loss factor of the subsystem, and the description ''effective'' is used to highlight

the fact that whereas the subsystem modal density, natural frequency statistics, and $P_{\text{in},k}$ may be insensitive to coupling to the rest of the system, this is not true of the loss factor. Lyon and DeJong (1995) suggest that the effective *in situ* loss factor should be computed as

$$\eta'_k = 1/(\omega n_k D_{0,kk}^{-1}). \tag{35}$$

This result is arrived at by writing the modal energy of the subsystem caused by a direct power input (which is $D_{0,kk}^{-1}P_{\text{in},k}$) as $P_{\text{in},k}/(\omega\eta'_k n_k)$, so that a "single subsystem" model employing η'_k yields the same result as a full solution of the SEA equations. Thus m'_k and B'_k are given by

$$m'_k = \omega\eta'_k n_k, \tag{36}$$

$$B'_k = \frac{\Delta}{\omega\eta'_k}. \tag{37}$$

Finally, in this section it can be noted that in deriving Eq. (33) the power inputs to the various subsystems have been assumed to be uncorrelated. Given that the power input to subsystem k is taken to be dominated by the local statistical properties of subsystem k, the correlation assumption will be valid if the subsystems have independent statistical properties. This is reasonable, given that the subsystem statistical variations will be caused by local uncertainties in such items as material properties, geometry, and nonstructural attachments (such as payload). Even where subsystems are part of the same physical element (for example, the in-plane and out-of-plane motions of a plate are usually modeled as separate subsystems), randomness will affect each subsystem differently, and uncorrelated power inputs can be expected.

C. The variance of the energy-power matrix

In order to investigate the term $\text{Var}(D_{\text{ran},ks})$ that appears in Eq. (33) it is helpful to rewrite Eq. (20) in the form

$$D_{kk}\overline{E}_k - \sum_{s\neq k}(-D_{ks}\overline{E}_s) = \overline{P}_k. \tag{38}$$

This equation can be viewed as a statement of power balance for subsystem k. The term on the right-hand side is the power input from external excitation, while the power dissipated through internal losses (given by $\omega\eta_k\overline{E}_k$) will form part of the first term on the left-hand side. The remaining contribution on the left-hand side will account for the power transferred to adjoining subsystems. With this interpretation, the term $-D_{ks}\overline{E}_s$ represents the power input to subsystem k arising from a prescribed energy $\overline{E}_s$ in subsystem s. The power input ultimately arises from forces acting on the boundary of subsystem k, and thus $-D_{ks}\overline{E}_s$ can be viewed as the power input caused by boundary forces that are consistent with the prescribed energy level $\overline{E}_s$. This argument then leads to the conclusion that $-D_{ks}$ is the power input arising from boundary forces that are consistent with a unit prescribed energy in subsystem s. Now the power input to subsystem k in response to prescribed boundary forces will have the form of Eq. (A1), with the terms ω_n representing the natural frequencies of the subsystem. Hence the relative variance of the term D_{ks} is given by

$$\text{Var}(D_{\text{ran},ks}) = D_{0,ks}^2 r^2(\alpha_{ks}, m'_k, B'_k), \tag{39}$$

where the function r^2 is given by Eqs. (A2) and (A6), and Eqs. (35)–(37) apply. The term α_{ks} is governed by the nature of the coupling between subsystems k and s and this is discussed in detail in the following section. Regarding the assumption made in Eq. (33) that the entries of $\mathbf{D}_{\text{ran}}$ are uncorrelated, it is clear from the above argument that $D_{\text{ran},ks}$ and $D_{\text{ran},jr}$ will be uncorrelated for $k\neq j$ since the statistical properties of different subsystems (k and j) are involved (this is analogous to the argument of the previous subsection concerning the correlation of the power inputs). It is less obvious that the terms will be uncorrelated for $k=j$, but it can be noted that for $s\neq r$ the relevant boundary forces differ, although the receiving subsystem is the same. If each term is expressed in terms of Eq. (A1), it can be anticipated that the numerator terms a_n for $D_{\text{ran},ks}$ will be uncorrelated with those for $D_{\text{ran},kr}$, thus justifying the assumption that the entries of $\mathbf{D}_{\text{ran}}$ are uncorrelated. A similar argument can be used to justify the uncorrelated nature of pairs of entries from $\mathbf{P}_{\text{in}}$ and $\mathbf{D}_{\text{ran}}$.

Equation (39) can be compared with an expression derived by Lyon and DeJong (1995) for the variance in the coupling loss factor η_{ks}, which should in principle have the same form as the present result. The Lyon and DeJong (1995) result is given by Eq. (12.3.7) of that reference, and actually differs fundamentally from Eq. (39): in Eq. (12.3.7) the variance is a symmetric function of the modal overlap factors of subsystems k and s, whereas Eq. (39) depends only on the modal overlap factor of subsystem k. Given this difference, it is helpful to compare both analyses with the exact form of the energy-power matrix given in Sec. II B. In particular, for the case of rain-on-the-roof excitation it follows from Eqs. (13)–(18) and Eq. (20) that

$$\mathbf{D} = (\tfrac{1}{2})\lceil q\rfloor\mathbf{M}_{\text{sym}}^{-1}\lceil\rho^{-1}n\rfloor \Rightarrow D_{ks} = (\tfrac{1}{2})q_k M_{\text{sym},ks}^{-1}\rho_s^{-1}n_s. \tag{40}$$

The unsymmetric nature of the random component of D_{ks} is highlighted by Eq. (40), since q_k is random (with an average value proportional to $\rho_k^{-1}n_k$) while ρ_s and n_s are deterministic. Although Eq. (40) is limited to the special case of rain-on-the-roof excitation, the form of the equation supports the general arguments given above regarding the structure of $\text{Var}(D_{\text{ran},ks})$.

D. The coupling coefficients α_{ks}

The coupling coefficient α_{ks} is determined by the nature of the force applied to subsystem k by subsystem s. As detailed by Langley and Brown (2003a, b) and outlined in the Appendix, if the generalized force applied to the nth mode of subsystem k is written as f_n, then

$$\alpha_{ks} = \text{E}[|f_n|^4]/\text{E}[|f_n|^2]^2. \tag{41}$$

This result will be independent of n if the generalized forces have common statistical properties. A number of cases are considered in what follows.

1. Multiple point couplings

If the two subsystems are coupled at N_p points, then the generalized force will have the form

$$f_n = \sum_{r=1}^{N_p} V_r \phi_n(x_r), \tag{42}$$

where $\phi_n(x_r)$ is the value of the nth mode shape evaluated at the rth coupling point x_r, and V_r is the complex amplitude of the force applied at x_r. The term V_r is a random complex number that will depend upon the dynamic response of the two coupled subsystems. It is reasonable to assume that many random factors influence V_r and by analogy with Schroeder's analysis of the random pressure in a room (Schroeder, 1962), V_r can be taken to be a zero mean complex Gaussian random variable, with uncorrelated real and imaginary parts of equal variance. Furthermore, V_r can reasonably be assumed to depend weakly on any one particular mode $\phi_n(x_r)$, and thus these two items can be considered to be approximately uncorrelated. It can then be shown from Eq. (42) that

$$\alpha_{ks} = \frac{(2K-2)}{N_p} + 2, \tag{43}$$

where K is given by Eq. (A8).

2. A line or area connection

If subsystems k and s are two-dimensional elements that are coupled along a line, then the generalized force will have the form

$$f_n = \int_L V(x)\,\phi_n(x)\,\mathrm{d}x, \tag{44}$$

where the integral is taken along the line junction and $V(x)$ is the force per unit length. Considerations based on the central limit theorem suggest that in this case f_n will tend to become complex Gaussian, as was assumed for V_r in the previous section. Thus

$$\alpha_{ks} = 2. \tag{45}$$

The same argument will hold for an area connection between two acoustic volumes, or between a plate and an acoustic volume.

E. Summary of the variance prediction process

The variance of the energy in subsystem j is given by Eq. (33), in which $\mathbf{D}_0$ is the standard SEA matrix and $\hat{E}_s$ is the ensemble averaged modal energy predicted by SEA. The variance terms that appear on the right-hand side of the equation are given by Eqs. (34) and (39), where the parameters m_k' and B_k' are given by Eqs. (35)–(37). The quantities α_k and α_{ks} are the only "non-SEA" parameters that appear in the analysis, and they describe the nature of the applied loading on subsystem k and the nature of the coupling between subsystems k and s. The value of α_k is given in the Appendix for various types of loading, while the value of α_{ks} will generally be 2, apart from the case of point connections, where Eq. (43) applies.

As noted in Sec. II C, it has been assumed in the forgoing derivations that the subsystem loss factors are deterministic. If the loss factors are taken instead to be random variables, then a Baysian approach can be used to calculate modified results for the variance: (*i*) the mean and mean squared values of the energy given by the present analysis for a specified set of loss factors can be considered to be conditional values; (*ii*) unconditional values for the mean and mean squared energy can be found by integrating the present results over the loss factor probability density function; and (*iii*) the unconditional variance can be found from the unconditional mean and mean squared values. The same approach could be used if the subsystem modal densities were taken to be random, although it should be noted that the modal density is by definition an ensemble averaged quantity. Randomness in this variable would imply gross changes to a subsystem beyond those expected in the normal ensemble, perhaps caused by very large mass changes or large changes in the system dimensions.

IV. NUMERICAL SIMULATIONS

A. General methodology

In order to validate the forgoing theory, numerical simulations have been performed for a range of plate and plate-acoustic systems. For each example system the Lagrange–Rayleigh–Ritz method has been used to generate the equations of motion, which are then solved for each member of a randomized ensemble. Each (rectangular) plate has been modeled by using the simply supported mode shapes as basis functions (considering thin plate theory and out-of-plane motions only), and a sufficient number of degrees-of-freedom have been employed to ensure convergence over the considered frequency range. Randomization of each plate has been achieved by adding ten point masses in random locations, with each mass having 1.5% of the total mass of the plate. Three rotational springs have also been added in random positions on each plate edge, to enhance the randomization at low frequencies. The degree of randomization is such that each plate has statistical overlap beyond the first few modes. In the coupled plate examples, coupling between two plates is enforced either by attaching a stiff linear spring between the out-of-plane displacement of a point in each plate, or by imposing edge coupling via stiff rotational springs that enforce rotational compatibility along a common edge (effectively the plates are then coupled through a simple support). In all cases aluminum plates are considered, with Young's modulus 7.2×10^{10} N/m^2, density 2800 kg/m^3, and Poisson ratio 0.3.

When an acoustic cavity is considered, the density of air is taken to be 1.3 kg/m^3, the speed of sound is 340 m/s, and a rectangular geometry is adopted. The system is modeled by using hard-walled modes, and randomness is introduced by randomly locating 20 cubic air pockets of approximate side length 0.1 m in which the speed of sound is artificially set at 34 m/s. The total volume of these pockets equates to 1% of the cavity volume.

Unless otherwise stated, an ensemble of 60 systems has been considered for each example, and the random locations

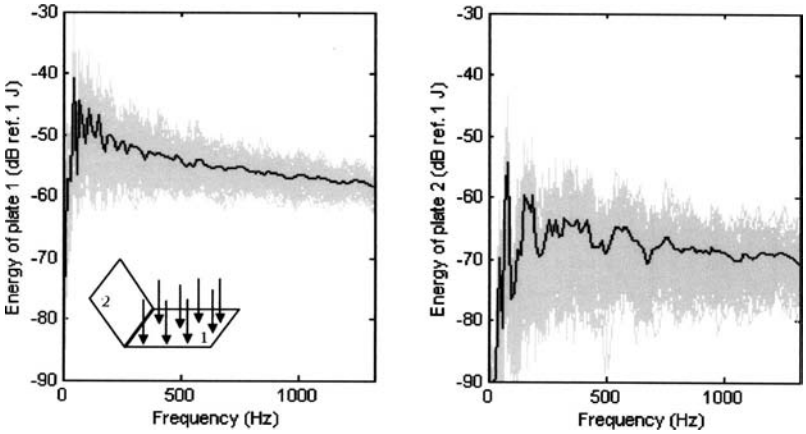


FIG. 1. Energy of plates 1 and 2 versus frequency, for two edge-coupled plates with rain-on-the-roof loading on plate 1. Gray: energy for each sample; Black: mean of energy over all samples.

of the masses/springs/air pockets have been selected from a uniform distribution over the possible geometric range.

B. Two coupled plates

The first example concerning two coupled plates is shown schematically in Fig. 1. The bending (out-of-plane) motions of two edge-coupled plates are considered, and rain-on-the-roof forcing is applied to plate 1 ($\alpha_1=2$). The properties of the plates are given in Table I, including the value of the modal overlap factor at 1 kHz; it can be noted that the modal overlap factor is a linear function of frequency, so the value at frequencies other than 1 kHz can readily be deduced. The energies of the two plates are shown in Fig. 1 as a function of frequency for each of the 60 generated samples, and the ensemble mean values are also shown as a function of frequency. It is clear that the mean energy is lower for plate 2 than for plate 1 while the scatter in the energy, i.e., the variance, is higher. At very low frequencies the variance is low for both plates, and the mean values show strong fluctuations with frequency indicating that the system has low statistical overlap to the point of being almost deterministic. At higher frequencies the mean curves become smoother, and the fluctuations are predominantly due to a lack of convergence arising from the limited number of samples employed in the ensemble. Similar fluctuations arise in the simulated results for the variance that are reported in what follows. The frequency range shown covers approximately 125 resonant modes of plate 1 and 33 resonant modes of plate 2.

TABLE I. Details of the plates considered in the numerical simulations.

Plate	Thickness (mm)	Size (m×m)	Loss factor η	Modal density (modes/Hz)	Modal overlap m at 1 kHz
1	5	1.35×1.2	0.02	0.0942	1.88
1b	5	1.35×1.2	0.03	0.0942	2.83
2	15	1.05×1.2	0.02	0.0245	0.49
3	5	1.05×1.2	0.02	0.0733	1.47
4	5	1.0×1.3	0.02	0.0756	1.51
4b	5	1.0×1.3	0.01	0.0756	0.76
5	5	1.05×1.1	0.03	0.0672	2.02
6	5	0.9×1.35	0.01	0.0707	0.71

Theoretical predictions for the mean (as given by SEA) and the relative variance (as given by the present theory) are compared with the simulation results in Fig. 2, where very good agreement is demonstrated. As in all the following figures, the simulation results are fluctuating curves and the theoretical predictions are smooth curves. The gray bands around the simulation curves are estimates of the 95% confidence intervals based on the finite number of samples employed (Bendat and Piersol, 1971). These bands are not reproduced on subsequent figures for the sake of clarity, but they give an indication of the level of simulation uncertainties in all of the following examples. At very low frequencies the SEA prediction becomes inaccurate due to strong coupling effects (Mace, 1994), while the present theory for the variance is further affected by a lack of statistical overlap in the simulations, so that the system is more deterministic than assumed by the theory. To put the level of relative variance shown in Fig. 2 into a practical context, it can be noted that a relative variance of 0.1 corresponds to a relative standard deviation of 0.32, meaning that the ''three sigma'' level ($1+3\times0.32$ times the mean) lies approximately 3 dB above the mean value on a log scale. The corresponding result for a relative variance of 1 is 6 dB, and for a relative variance of 10 the result is 10 dB.

The effect of replacing the rain-on-the-roof forcing with single point forcing ($\alpha_1=2.75$) is demonstrated by the results shown in Fig. 3. The mean values of the energy are more or less unchanged, while there is an increase in the relative variance of both plates that is well predicted by the theory. The effect of changing the nature of the coupling is shown in Fig. 4, where the line junction between the plates ($\alpha_{12}=2$) is replaced by three point connections [$\alpha_{12}=3.17$, from Eq. (43)]. The relative variance of the driven plate is relatively unchanged, while the relative variance of plate 2 is increased due to an increase in the effective uncertainty in the coupling, as measured by α_{12}.

In considering the accuracy or otherwise of the variance predictions shown in Figs. 2–4 it should be borne in mind that SEA is an approximate theory. The present analysis builds upon SEA and is therefore inherently approximate. To consider Fig. 4, for example, the level of agreement between the predicted mean energy and the simulations is not exact, and the discrepancies shown are fairly typical of SEA. At low frequencies the basic hypotheses employed in SEA (for

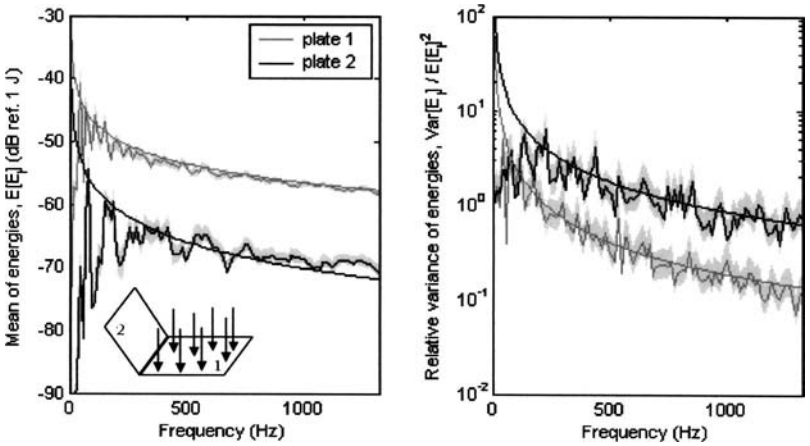


FIG. 2. Mean and relative variance of energy of plates 1 and 2 versus frequency, for two edge-coupled plates with rain-on-the-roof loading on plate 1. The 95% confidence intervals resulting from the finite number of samples in the random ensemble are shown in light gray.

example weak coupling) break down and a good prediction cannot be expected. The variance predictions are further affected by a possible lack of applicability of the GOE at low frequencies. The relative variance predictions shown in Fig. 4 are far from perfect for plate 2 below 600 Hz, although the prediction can still be considered to be good overall. A key point is that for this problem SEA requires the solution of a trivial 2×2 matrix problem, while the conventional analysis requires many degrees of freedom and the consideration of a large ensemble.

C. Plate-acoustic system

This system is shown schematically in Fig. 5. An air filled cavity of dimension $1.45\times 1.35\times 1.2$ m^3 has rigid walls apart from one flexible wall which consists of plate 1b in Table I. The acoustic loss factor is set at $\eta=0.03$, giving a modal overlap factor of $m=2.8$ at 500 Hz and $m=0.7$ at 250 Hz, this result being approximately proportional to the square of the frequency (Lyon and DeJong, 1995). Approximately 35 modes of the cavity fall below 500 Hz. The response to a point acoustic source is shown in Fig. 5, where it can be seen that there is good agreement between the simulations and the theory above 150 Hz. Below that frequency there are very few acoustic modes in the cavity and the degree of statistical overlap is low; hence the cavity is near deterministic, and the variance is much lower than the predicted value. Results for the case of a point load applied to the plate are shown in Fig. 6, where again there is good agreement between the simulations and the theory beyond 150 Hz.

Results for the mean and variance of the frequency band averaged energy are shown in Fig. 7 for the case of a point load applied to the plate. A moving frequency average has been taken on each simulated sample, and the bandwidth $\Delta=5\omega\eta$ has been employed. By comparing Fig. 7 with Fig. 8, it can be seen that band averaging leads to a significant reduction in the relative variance of both the plate and the acoustic volume, and this reduction is well predicted by the present theory.

Throughout Figs. 6–8 the simulation results for the relative variance fluctuate significantly with frequency; this is partly due to the limited number of simulations employed and partly due to difficulties in randomizing the cavity in a homogeneous way within the context of the present simple analytical model.

D. Three- and four-plate networks

A schematic of a three-plate network is shown in Fig. 8, where the plate numbering corresponds to the properties given in Table I. Plates 1 and 3 are coupled by a line junction while plates 3 and 4 are coupled at three points. In this case 120 samples rather than 60 samples have been considered in the random ensemble, so that the simulated relative variance curves are smoother and more easily distinguishable in the figure. It can be seen that the relative variance is very well

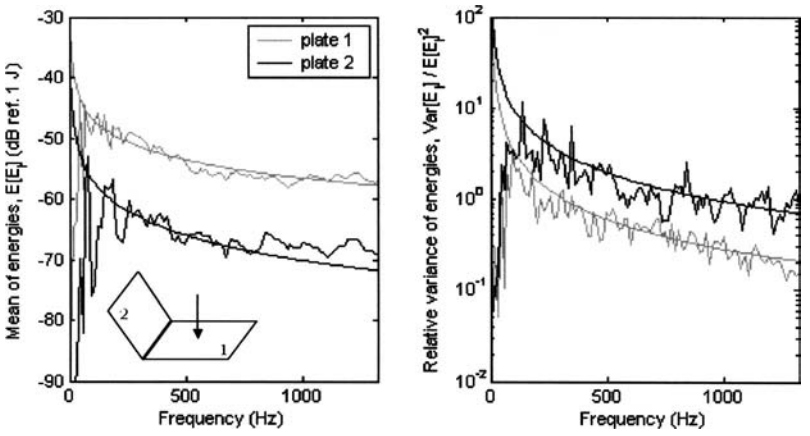


FIG. 3. Mean and relative variance of energy of plates 1 and 2 versus frequency, for two edge-coupled plates with point loading on plate 1.

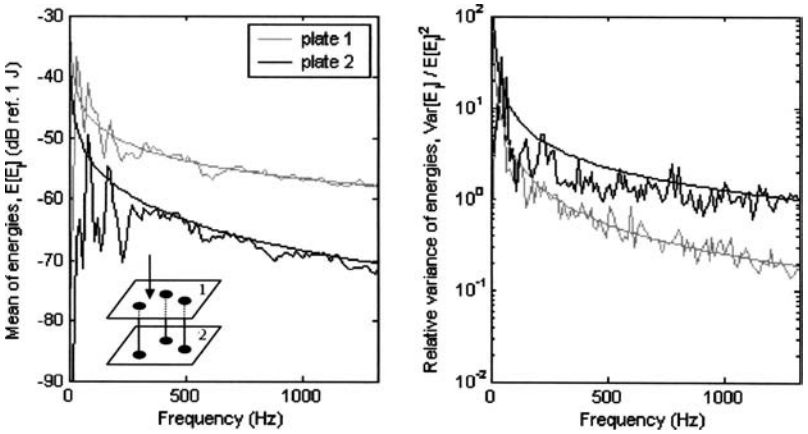


FIG. 4. Mean and relative variance of energy of plates 1 and 2 versus frequency, for two point-coupled plates with point loading on plate 1.

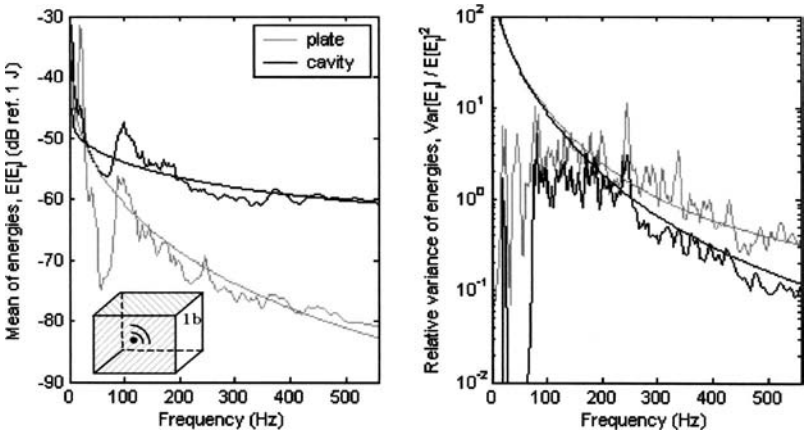


FIG. 5. Mean and relative variance of the energy of the plate and cavity versus frequency, with an acoustic point source in the cavity.

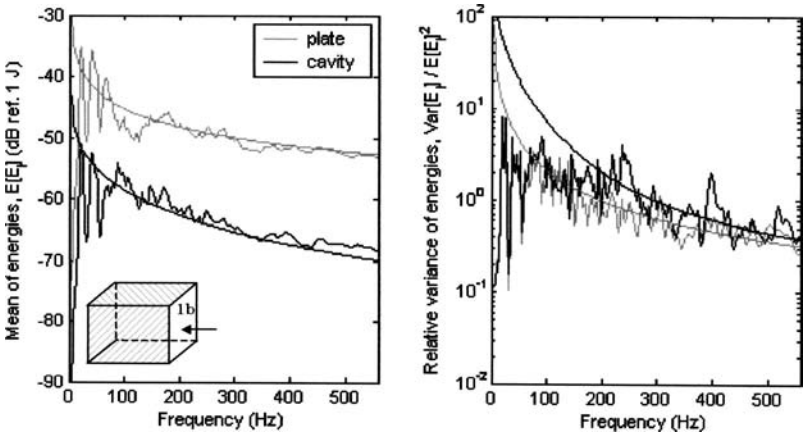


FIG. 6. Mean and relative variance of the energy of the plate and cavity versus frequency, with a point loading on the plate.

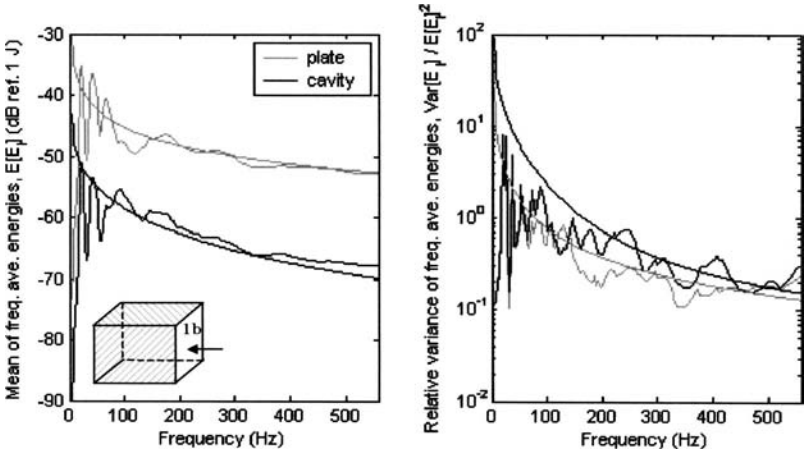


FIG. 7. Mean and relative variance of the frequency-band averaged energy of the plate and cavity versus frequency, with a point loading on the plate. The moving average is performed over the band $\Delta = 0.15\omega$.

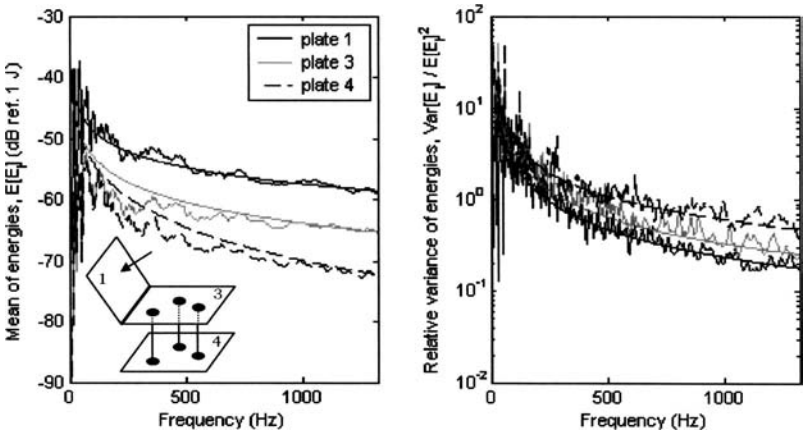


FIG. 8. Mean and relative variance of energy of plates 1, 3, and 4 versus frequency, for three plates coupled by edge and points with point loading on plate 1.

predicted by the present method. The variance increases with distance from the driven subsystem (plate 1), as might be expected on physical grounds, since power must be transferred through an increasing number of random subsystems.

The final example is the four-plate network shown in Fig. 9, with the plate properties being listed in Table I. All of the previous examples have employed a constant loss factor throughout the structure, so that the system has stiffness proportional damping. This has been done for computational convenience, since proportional damping allows the system equations of motion to be uncoupled when expressed in modal coordinates, thus producing a large reduction in the computational effort required for the random simulations. There may be a concern, however, that proportional damping is a special case, and thus the final example has been selected to have highly nonproportional damping. The results shown in Fig. 9 indicate that the relative variance is well predicted by the present theory. It can be noted that although plates 5 and 6 are both directly connected to the excited plate, the relative variance differs due to the difference in the loss factors.

In Fig. 10 results are shown for the relative variance predicted by the analysis of Lyon and DeJong (1995) for the final four-plate example. Although the analysis predicts the correct trends, there is a lower level of quantitative agreement with the simulations than that displayed in Fig. 9. The present analysis differs from that of Lyon and DeJong (1995) in that (*i*) the GOE is employed, rather than the assumption of Poisson natural frequencies; (*ii*) the random part of the energy-power matrix **D** is not forced to be symmetric; (*iii*) the constraint represented by Eq. (31) is applied to the energy-power matrix; and (*iv*) the relative variance of the matrix entry D_{ks} is taken to depend only on the properties of subsystem k via Eq. (39), which contrasts with Eq. (12.3.7) of Lyon and DeJong (1995). Clearly these are important issues and they have a significant effect on the predicted variance.

V. CONCLUDING REMARKS

A method has been presented for predicting the variance of the energy levels in a general built-up system within the context of SEA. The key equations behind the method are summarized in Sec. III E, and these results have been validated by application to a range of examples in Sec. IV. The key assumptions underpinning the method are (*i*) SEA provides a good estimate of the mean subsystem energies, so that the normal requirements of SEA (for example weak coupling) are met; (*ii*) each subsystem is sufficiently random to exhibit statistical overlap, so that GOE statistics apply to the natural frequencies. Condition (*i*) implies that the current analysis is not intended to correct any bias in the prediction of the mean energy arising, for example, from strong coupling (Mace, 1994) or wave filtering (Langley, 1994). If condition (*ii*) does not apply, then the statistics of the response will be sensitive to details in the nature of the system uncertainty and a much more complicated analysis will be required—in fact, the problem may actually be intractable in this case. Fortunately, previous studies have shown that rela-

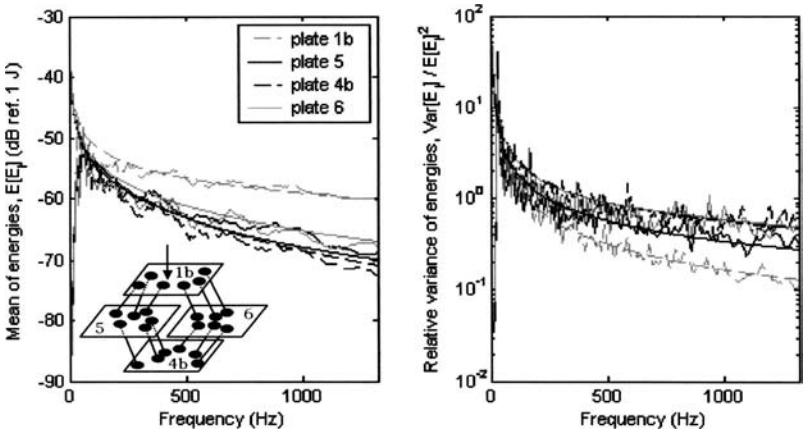


FIG. 9. Mean and relative variance of energy of plates 1b, 5, 4b, and 6 versus frequency, for four point-coupled plates with point loading on plate 1.

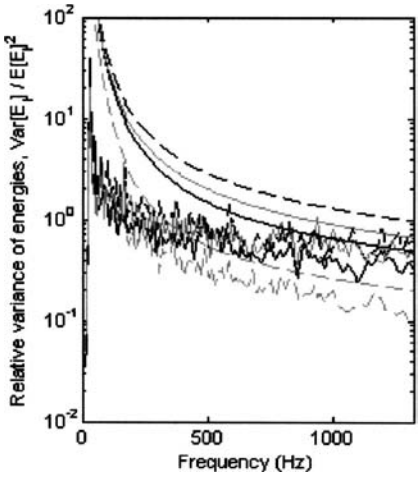


FIG. 10. Relative variance of energy of plates 1b, 5, 4b, and 6 versus frequency, for four point-coupled plates with point loading on plate 1: prediction according to the formulation from Lyon and DeJong (1995).

tively little randomness is required to achieve GOE statistics (see, for example, Weaver, 1989; Mehta, 1991; Lobkis *et al.*, 2000), although there is scope for further work on quantifying at the design/analysis stage whether condition (*ii*) will be met for a particular structure. Further work is also required to establish the statistical distribution of the energy, so that confidence bands can be established; there is some evidence to suggest that the log-normal distribution is appropriate (Fahy and Mohammed, 1992, Manohar and Keane, 1994, Lyon and DeJong, 1995), in which case the distribution would be fully defined by the mean and variance.

The variance expressions derived here are very straightforward to apply, and they could readily form a rapid post-processing step following the application of SEA to a large complex system.

ACKNOWLEDGMENTS

This work was sponsored by the Air Force Research Laboratory, Space Vehicles Directorate, Kirtland AFB, NM, under SBIR Phase II Contract No. F29601-02-C-0109.

APPENDIX: THE VARIANCE OF A MODAL SUM

Langley and Brown (2003a, b) have considered the ensemble variance of a quantity in the form

$$T=\frac{1}{\Delta}\int_{\omega-\Delta/2}^{\omega+\Delta/2}\sum_n \frac{a_n}{(\omega^2-\omega_n^2)^2+(\eta\omega_n)^2}\,\mathrm{d}\omega. \tag{A1}$$

This expression arises when considering the energy of the response of a single dynamic subsystem, such as a beam, plate, or acoustic volume: the terms ω_n are the natural frequencies, η is the loss factor, and the terms a_n depend upon the nature of the loading. For a random system, both ω_n and a_n vary over the ensemble. By assuming that the natural frequencies have statistics consistent with the Gaussian orthogonal ensemble (GOE) in random matrix theory, Langley and Brown (2003a, b) have shown that the relative variance of T (i.e., the variance divided by the square of the mean) is given by

$$r^2(\alpha,m,B)=\frac{\alpha-1}{\pi m}\left(\frac{1}{B^2}\right)\left\{2B\left[\frac{\pi}{2}-\tan^{-1}\left(\frac{1}{B}\right)\right]-\ln(1+B^2)\right\}+\frac{1}{(\pi m)^2}\left(\frac{1}{B^2}\right)\ln(1+B^2), \tag{A2}$$

where

$$\alpha=\frac{\mathrm{E}[a_n^2]}{\mathrm{E}[a_n]^2}, \tag{A3}$$

$$m=\omega\eta n, \tag{A4}$$

$$B=\frac{\Delta}{\omega\eta}, \tag{A5}$$

and n is the modal density of the system. Note that the parameter m is the modal overlap factor of the system. Equation (A2) is valid for intermediate to large values of m (typically $\pi m>3$); an alternative result that is valid for all m is available for the case $B=0$ (Langley and Brown, 2003a), and this has the form

$$r^2(\alpha,m,0)=\frac{1}{\pi m}\left\{\alpha-1+\frac{1}{2\pi m}[1-\exp(-2\pi m)]+\mathrm{E}_1(\pi m)\left[\cosh(\pi m)-\frac{1}{\pi m}\sinh(\pi m)\right]\right\}, \tag{A6}$$

where E_1 is the exponential integral (Abramowitz and Stegun, 1964).

If the system is subjected to N_p incoherent point loads, then the appropriate value of α is

$$\alpha=\frac{(K-2)}{N_p}+2, \tag{A7}$$

$$K=\frac{\mathrm{E}[\phi_n^4(x)]}{\mathrm{E}[\phi_n^2(x)]^2}, \tag{A8}$$

where $\phi_n(x)$ is the mode shape at some location x. For Gaussian mode shapes, $K=3$, although numerical studies have shown that $K=2.75$ is more appropriate for plate structures (Weaver, 1989; Langley and Brown, 2003a). If the energy response of the system to a single point force is considered, and this response is then averaged over N_p point locations, then Eq. (A7) is modified to

$$\alpha=\frac{(K-1)}{N_p}+1. \tag{A9}$$

The value of α appropriate to other types of loading can similarly be derived. One other case worthy of note arises when the generalized force in each mode is a zero mean complex Gaussian variable, with uncorrelated real and imaginary parts of equal variance: in this case $\alpha=2$.

[1]Note that a pseudo-inversion of Eq. (9) is not acceptable here: for specified **T** this will yield the force spectrum of minimum norm, but this will not necessarily correspond to the true force spectrum **S**. Hence the power **P** will be incorrectly computed by substituting this result for **S** into Eq. (10) to yield **P** in terms of **T** in the form of Eq. (7).

Abramowitz, M., and Stegun, I. A. (**1964**). *Handbook of Mathematical Functions* (Dover, New York).
Bendat, J. S., and Piersol, A. G. (**1971**). *Random Data: Analysis and Measurement Procedures* (Wiley-Interscience, New York), pp. 110–114.
Bertelsen, P., Ellegaard, C., and Hugues, E. (**2000**). ''Distribution of eigenfrequencies for vibrating plates,'' Eur. Phys. J. B **15**, 87–96.
Craik, R. J. M. (**1996**). *Sound Transmission Through Buildings: Using Statistical Energy Analysis* (Gower, Aldershot).
Davy, J. L. (**1986**). ''The ensemble variance of random noise in a reverberation room,'' J. Power Sources **107**, 361–373.
Ellegaard, C., Schaat, K., and Bertelsen, P. (**2001**). ''Acoustic chaos,'' Phys. Scr. **90**, 223–230.
Fahy, F. J., and Mohammed, A. D. (**1992**). ''A study of uncertainty in applications of SEA to coupled beam and plate systems. I Computational experiments,'' J. Sound Vib. **158**, 45–67.
Fredö, C. R. (**1997**). ''SEA-like approach for the derivation of energy flow coefficients with a finite element model,'' J. Sound Vib. **199**, 645–666.
Guyader, J. L., Boisson, C., and Lesueur, C. (**1982**). ''Energy transmission in finite coupled plate. Part I: Theory,'' J. Sound Vib. **81**, 81–92.
Keane, A. J., and Price, W. G. (**1987**). ''Statistical energy analysis of strongly coupled systems,'' J. Sound Vib. **117**, 363–386.
Langley, R. S. (**1989**). ''A general derivation of the statistical energy analysis equations for coupled dynamic systems,'' J. Sound Vib. **135**, 499–508.
Langley, R. S. (**1990**). ''A derivation of the coupling loss factors used in statistical energy analysis,'' J. Sound Vib. **141**, 207–219.
Langley, R. S., and Bercin, A. N. (**1994**). ''Wave intensity analysis of high frequency vibrations,'' Philos. Trans. R. Soc. London, Ser. A **346**, 489–499.
Langley, R. S., and Brown, A. W. M. (**2003a**). ''The ensemble statistics of the energy of a random system subjected to harmonic excitation,'' J. Sound Vib. (accepted for publication).
Langley, R. S., and Brown, A. W. M. (**2003b**). ''The ensemble statistics of the band averaged energy of a random system,'' J. Sound Vib. (accepted for publication).
Lobkis, O. I., Weaver, R. L., and Rozhkov, I. (**2000**). ''Power variances and decay curvature in a reverberant system,'' J. Sound Vib. **237**, 281–302.
Lyon, R. H. (**1969**). ''Statistical analysis of power injection and response in structures and rooms,'' J. Acoust. Soc. Am. **45**, 545–565.
Lyon, R. H. (**1975**). *Statistical Energy Analysis of Dynamical Systems: Theory and Applications* (MIT, Cambridge).
Lyon, R. H. (**2001**). ''Variability in SEA,'' on the CD-ROM: *Proceedings of India-USA Symposium on Emerging Trends in Vibration and Noise Engineering, The Ohio State University, Columbus, OH, December 10–12, 2001*, edited by A. Selamet, R. Singh, and B. C. Nakra.
Lyon, R. H., and DeJong, R. G. (**1995**). *Theory and Application of Statistical Energy Analysis* (Butterworth-Heinemann, Boston).
Mace, B. R. (**1994**). ''On the statistical energy analysis hypothesis of coupling power proportionality and some implications for its failure,'' J. Sound Vib. **178**, 95–112.
Mace, B. R., and Shorter, P. J. (**2000**). ''Energy flow models from finite element analysis,'' J. Sound Vib. **233**, 369–389.
Manohar, C. S., and Keane, A. J. (**1994**). ''Statistics of energy flows in spring-coupled one-dimensional systems,'' Philos. Trans. R. Soc. London **346**, 525–542.
Mehta, M. L. (**1991**). *Random Matrices* (Academic, London).
Schroeder, M. R. (**1962**). ''Frequency-correlation functions of frequency responses in rooms,'' J. Acoust. Soc. Am. **34**, 1819–1823.
Soize, C. (**2003**). ''Random matrix theory and non-parametric model of random uncertainties in vibration analysis,'' J. Sound Vib. **263**, 893–916.
Weaver, R. L. (**1989**). ''Spectral statistics in elastodynamics,'' J. Acoust. Soc. Am. **85**, 1005–1013.
Woodhouse, J. (**1981**). ''An approach to the theoretical background of statistical energy analysis applied to structural vibration,'' J. Acoust. Soc. Am. **69**, 1695–1709.

Longitudinal surveys on effects of changes in road traffic noise—annoyance, activity disturbances, and psycho-social well-being

Evy Öhrström[a)]
Department of Environmental Medicine, Sahlgrenska Academy at Göteborg University, Box 414, SE 405 30 Göteborg, Sweden

(Received 18 January 2003; revised 31 October 2003; accepted 4 November 2003)

The adverse effects of long-term exposure to a high volume of road traffic were studied in socio-acoustic surveys in 1997 and in 1999 after a substantial reduction in road traffic. The results obtained in 1997 showed a similar response pattern as in previously performed studies in the area in 1986 [Öhrström, J. Sound Vib. **122**, 277–290 (1989)]. In 1999, road traffic had been reduced from 25 000 to 2400 vehicles per day, and this resulted not only in a large decrease in annoyance and activity disturbances, but also in a better general well-being. The results suggest that a reduction in both noise and other pollutants from road traffic contribute to these effects. To be able to use the outdoor environment and to have the possibility to keep windows open is essential for general well-being and daily behavior, which implies that access both to quiet indoor and outdoor sections of the residency is of importance for achievement of a healthy sound environment. More knowledge of long-term health consequences of exposure to noise and simultaneous pollutants from road traffic is needed. Studies should focus more on ''softer'' health outcomes and well-being than hitherto and preferably be performed in connection with traffic abatement measures. [DOI: 10.1121/1.1639333]

PACS numbers: 43.50.Qp [DKW] Pages: 719–729

I. INTRODUCTION AND AIM

There is substantial evidence in the literature that environmental noise causes a variety of adverse health effects and the evidence is strongest for annoyance, speech interference, acute sleep disturbances, and children's performance in school (IEH, 1997; WHO, 2000). Knowledge of long-term health effects such as effects on psychological and physiological well-being is very limited (Health Council of the Netherlands, 1994; Passchier-Vermeer, 1996; Lercher 1996). The Health Council of the Netherlands and Passchier-Vermeer classified the evidence for a causal relationship between possible long-term effects and environmental noise. The authors concluded that there is sufficient evidence for hearing loss, hypertension, ischaemic heart disease, annoyance, and sleep disturbance (e.g., changes in sleep pattern, awakening, sleep stages, subjective sleep quality, heart rate, mood next day). For some presumed adverse health effects, the evidence for a causal relationship with environmental noise was considered weak and/or the number of studies too limited for reliable conclusions to be drawn. Examples of such effects were immune effects, birth weight, psychiatric disorders, and psycho-social well-being. A UK workshop on nonauditory effects of noise dealt with the effects on sleep, its methodology, and other research priorities (IEH, 1997). The general conclusion was to encourage longitudinal studies (i.e., repeated studies on the same population) and particularly concentrate on susceptible groups, as well as field experiments or intervention studies because these are most informative. However, laboratory studies were considered necessary for clarifying specific causal hypotheses and to define dose-effect relationships. Knowledge gained in laboratory studies (Öhrström and Björkman, 1988) shows that exposure to road traffic noise for nine nights during a period of two weeks negatively affected sleep quality, performance, and mood, and increased tiredness both in the morning and during the daytime. No habituation was seen for these effects. If people who have resided in areas exposed to high levels of road traffic noise for many years are affected in a similar way, they may develop long-term conditions, e.g., physiological and psychological symptoms. This can only be studied in longitudinal studies and/or intervention studies in connection with extensive noise abatement activities as ''natural experiments'' give vital information that cannot be obtained in cross-sectional studies.

The aim of the present longitudinal study was to assess various presumed long-term health effects—especially in terms of psychological and physiological well-being—of road traffic noise and the effects of changes in noise exposure on (1) annoyance, (2) activity disturbances indoors and outdoors, (3) sleep quality, (4) psychological and physiological well-being, and (5) perception and use of the living environment.

II. BACKGROUND TO THE STUDY

Västra Bräckevägen on the island of Hisingen in the city of Göteborg, Sweden, has had a very heavy traffic load for many years, with 25 000 to 30 000 vehicles per 24 hours and $L_{Aeq,24h}$-levels of more than 65 dB. To improve the environmental situation of people living along this road, a number of measures have been taken over the years. Socio-acoustic sur-

a)Electronic mail: evy.ohrstrom@envmed.gu.se

veys were made along Västra Bräckevägen, before and after these measures in 1986 and 1987 (Öhrström, 1989; Björkman *et al.*, 1988), to shed light on the occurrence of different effects of noise on the population and on whether the measures taken had had an effect on the experience of disturbance caused by noise on sleep quality and on perceived health and well-being. These investigations (Björkman *et al.*, 1988) showed that the measures enacted against noise (prohibition of heavy traffic during the night, porous asphalt, speed reductions, and traffic signs that showed the message ''You are driving too fast'') were not sufficient to reduce the noise and its adverse effects. The study (Öhrström, 1989) showed that 67% of the interviewed participants were very annoyed by road traffic noise in the exposed area as opposed to 19% in the control area in 1986. There were also significant differences for difficulties falling asleep and sleep quality reported in the morning and for perceived tiredness during the morning and during the day between the areas. The participants were also interviewed about the occurrence of a number of psychological and physiological symptoms chosen to reflect known effects due to sleep deprivation and suffering from stress. Significant differences were found for some of these reported psychological and physiological symptoms between the exposed and the control area. These symptoms were ''Feel very tired,'' ''Nervous stomach,'' ''Want to be left alone,'' and ''Headaches.'' Other symptoms asked for in the interview, ''Irritable,'' ''Depressed,'' and ''Anxious/nervous,'' were also more frequent in the exposed area but the difference was not significant between the areas. Although no firm conclusions can be drawn from just one study, it was hypothesized from the results of the study that high levels of road traffic noise during the day and night have long-term effects on psycho-social well-being in terms of psychological and physiological symptoms.

A very good opportunity to perform a new intervention study in the same areas turned up about 10 years later, in 1996, when plans to improve the living environment around Västra Bräckevägen and to facilitate road transport were decided upon. These measures were planned to be very extensive and included changes in the road system and construction of a new tunnel, the Lundby Tunnel, which was opened for traffic in January 1998.

III. METHOD AND MATERIAL

A. Investigation area

The areas investigated and the measurement locations are shown on the map in Fig. 1. The northern boundary of the area is a busy thoroughfare, Västra Bräckevägen, while the boundaries to the west, east, and south are smaller local streets. The investigation area was divided into an exposed (unbroken lines) and a control area (dotted lines) in which the houses were situated 25–67 m and 125–405 m, respectively, from the heavily trafficked main road.

The houses in the area are very similar, i.e., smaller, detached one-family houses built during the 1930s and surrounded by nice gardens of between 200 and 400 square meters. All houses were privately owned. The bedroom and living room windows in most of the houses in the exposed

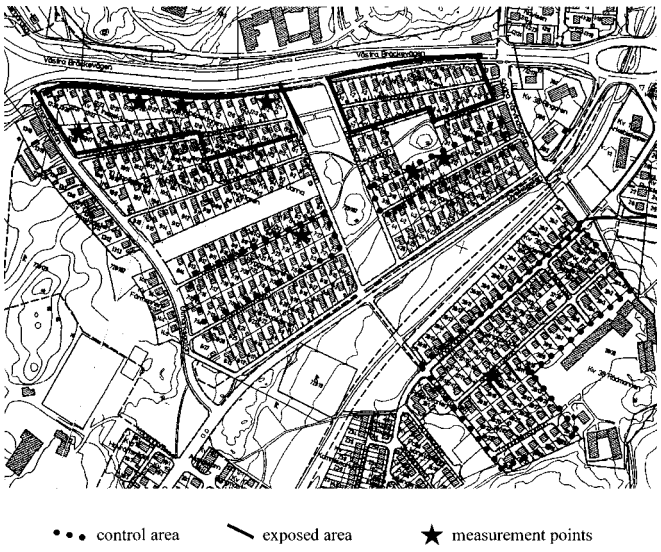


FIG. 1. Investigation area.

area do not face toward Västra Bräckevägen; only 16% of the houses' bedroom windows and 20% of the houses' living room windows faced toward the road. About 40% of the houses in the exposed area had double-paned windows in the living room, kitchen, and bedrooms while the others had some form of triple-paned windows. Thirty-five per cent of the bedrooms in the control area had double-paned windows and 43% had double-paned windows in the living room and kitchen. A higher percentage of residents ($p<0.02$) had added extra facade insulation in the exposed area (55% as compared to 34% in the control area).

B. Assessment of noise exposure

Noise exposure was determined by measurements and calculations. Information about the traffic load per annual mean 24-h weekday for day, night, and type of vehicle was obtained from the Traffic Office in Göteborg.

Calculations of equivalent sound levels $L_{Aeq,24h}$—The calculations were made according to the Nordic Prediction Method (Nordic Prediction Method, 1996). The input data for this model are: vehicles/24h, % heavy vehicles, speed limit, distance road-receiver, distance road-screen, receiver height, ground (soft/hard), start angle, and stop angle. The calculation program has a precision of ± 2 dB. Calculations of noise levels were made for the most severely noise-exposed side (north-western) and for the quieter side (south-eastern) at two heights, 2.5 and 5 m, which correspond to the first and second floors of the houses. The difference in the sound level between 2.5 and 5 m was marginal (± 1 dB).

Outdoor noise measurements—A sound level meter, Larson & Davis type LD 820, which was operated as a remote station via a wireless transmission system, was used. Measurements were made continuously for 3 to 4 days at eight different locations in the study area. A mean for each position was calculated for $L_{Aeq,24h}$, L_{01}, L_{90} and L_{Amax}, as well as noise events >70 dBA for three periods: 24 hours, daytime (06–22) and night time (22–06). The microphone was placed on a rod 1.5 m above ground level in a free-field

position, that is without reflection from the facade. The distance to the road was the same as to the facade that faced toward the thoroughfare.

Measurements of facade insulation—Measurements were made of the reduction at the facade ($R'45°,w$) in the bedrooms of seven of the houses. These measurements were made according to Swedish Standard SS 02 52 54 (1980) and ISO 140-5 with a Norsonic AS 840 (1988) measurement system. The evaluation was carried out according to SS EN ISO 717-1 (1996).

Determination of the equivalent sound levels inside and outside the living room and bedroom—The indoor levels were estimated from information in the surveys as to the type of windows and in which direction the windows faced, measured facade reduction ($R'45°,w$) and the calculated outdoor equivalent levels on each side of each house. Double-paned windows were given a reduction of 31 dBA and triple-paned windows a reduction of 34 dB according to the results of the facade reduction measurements, which varied between 29 and 38 dBA. Outdoor levels in $L_{Aeq,24h}$ outside the living room and bedroom were determined from information in the survey and calculations of outdoor ($L_{Aeq,24h}$) levels on each side of each house.

C. Assessment of effects

Effects of exposure to road traffic noise were evaluated by questionnaires, similar to the ones used in previous studies (Öhrström, 1989; Björkman *et al.* 1988). The questionnaire was sent with an introductory letter to all individuals selected for the study. The letter described the study as a general living environment investigation to avoid a focus on road traffic noise. The person was informed about how he or she had been selected, how the answers would be treated and presented, and that his or her participation was voluntary. Two reminder letters, at an interval of 10 days, were sent to those who failed to respond. The questionnaire used in 1997 contained a total of 41 questions with subquestions and consisted of five parts: (I) the home and the home environment; (II) sleep and sleeping environment; (III) health and well-being; (IV) general questions; and (V) space for the respondents' own comments.

The first part contained questions about length of residence, type of window, the position of bedroom and living room windows in relation to the roads, and satisfaction with the dwelling and the nearby surroundings. This section also contained questions about sources of annoyance in the environment, such as noise from road traffic, neighbors, aircraft, and industries, exhaust fumes and vibrations from motor vehicles, dust or soot in the air, and odors from industries. The response categories to these annoyance questions were ''not at all,'' ''not very,'' ''rather,'' and ''very.'' Some of these annoyance questions (exhaust fumes, vibrations, and noise) were complemented with numerical response scales from 0 to 10 with endpoint markings ''not at all annoyed'' and ''extremely annoyed.'' These numerical scales were used to facilitate comparisons between international studies. The formulation of this annoyance question (and also a 5-graded verbal category scale) is the outcome of several years of discussions and work among researchers in the field of effects of community response to noise and has recently been included in an ISO specification (Technical Specification ISO/TS15666, 2003). Other questions concerned activity disturbances indoors and outdoors due to road traffic with the response categories ''no'' or ''yes'' and if ''yes,'' ''How disturbing is this?:'' ''not very,'' ''rather,'' and ''very.'' Questions on activity disturbances indoors concerned conversation, radio/TV, concentration, rest/relaxation, difficulties in falling asleep, awakenings and not keeping windows open as often as desired. Questions on outdoor activity disturbances were posed on conversation, relaxation, and ability to be on the terrace or patio. *The second part* of the questionnaire contained 15 questions on sleep and sleep environment, e.g., the location of the bedroom, time for going to bed and getting out of bed, and questions on the presleep period, the sleep period and the morning. *The third part* of the questionnaire contained questions about general health problems and the respondents also answered how often [(1) seldom/never, (2) a few times a month, (3) a few times a week, or (4) every day] they experienced various symptoms or felt: very tired, irritable/sullen, unhappy/depressed, anxious/nervous, unsociable/wants to be left alone, or had headaches or an uncomfortable/upset stomach. A total score for psycho-social well-being was calculated from the sum of the seven symptoms. *The fourth part* of the questionnaire dealt with general questions on civil status, age, sex, employment, and sensitivity to noise, sensitivity to cold/heat and sensitivity to dust/soot in the air. The question on noise sensitivity was phrased ''How would you in general describe your sensitivity to noise?'' The response categories were (1) not at all sensitive, (2) not very sensitive, (3) rather sensitive, and (4) very sensitive.

In the follow-up study in 1999 the same questions were posed as in 1997. *In addition* to these questions, different questions were used to obtain knowledge of the participants' perception of how the building of Lundby tunnel and other changes in the road system affected them. These comparative questions had three or five response categories and concerned: satisfaction with the home, the home environment, notice noise, disturbances of a number of activities, sleep and sleep habits, and psycho-social well-being. The reason for asking these questions was that both the verbal category response scales and the numerical response scales are rather limited. In long-term follow up studies like this, the comparison scales are complementary to the other scales and give the respondents a possibility to make their own comparison. Examples of how these comparative questions were phrased to collect information on sleep disturbances and sleep habits are: ''*If you compare with the situation before 1998, before the Lundby tunnel was built, how often do you...*?:'' sleep with windows open during the night (more seldom, no difference, more often), have difficulties in falling asleep (much more often, somewhat more often, no difference, somewhat less often, much less often), and wake up at night (less often, no difference, more often). In addition there are questions on sleep quality (worse, no difference, better) and tiredness in the morning (more tired, no difference, more alert).

Finally, open questions were used to catch the respondents' reactions to the changes in their environment that were

not possible to cover (or to be foreseen) by questions in the questionnaire. Three such open questions were asked: ''Do you think that the way you use your home and your home environment has been affected by the building of the Lundby Tunnel?;'' ''Do you think that your sleeping habits and your sleep has been affected by the building of the Lundby Tunnel?;'' and ''Do you think that your general well-being has been affected in some way by the building of the Lundby Tunnel?.'' The response categories were ''no'' or ''yes'' and, if ''yes,'' ''In what way?''

D. Study population

After determining the exposed area and control area, a geographical selection of the population was done using a population register obtained from the Town Planning Office in Göteborg. One or two individuals in each household aged between 18 and 80, who had lived in the area for at least one year, were chosen for the study. [The mean age of the respondents was similar in the two areas, 51.0 and 50.4, respectively, in the exposed and the control area. The length of residence was 11.2 (s.d. 10.7) years in the exposed and 13.3 (s.d. 13.2) years in the control area.] A total of 230 individuals in 140 households were included. There were 142 respondents (50 in the exposed area and 92 in the control area), which meant a response rate of 62%. These individuals represented 69% of the households in the area. In the follow-up study the selection base had decreased by 14 individuals and the number of respondents decreased from 142 to 120 (45 in the exposed area and 75 in the control area), which meant a response rate of 85% in 1999 of those who participated in 1997.

The first part of the investigation was carried out between October and December of 1997. The follow-up study, carried out one year after the tunnel was opened for road traffic, was conducted in April and May of 1999.

E. Statistical analysis and treatment of data

The X^2 test, Mann-Witney U-test and T-test were used for tests of differences between different groups. The paired T-test and Wilcoxon sign rank test were used to test differences in results in the same group. The T-test for one group was used for analyses of results from the respondents' own comparison with the situation before the Lundby tunnel was built (e.g., ''less disturbed,'' ''no difference,'' and ''more disturbed'') of annoyance, activity disturbances, sleep and well-being, to obtain the mean value of a group. The covariation between different variables was investigated and tested with the Spearman rank correlation test, r_s. No significant change was expected in different effects studied in the control area between 1997 and 1999, and therefore the two-sided test was used. Based on the findings in the studies performed in 1986 and 1987 a significant change in adverse effects of noise including measures of psycho-social well-being was hypothesized in the exposed area in 1999. The one-sided test was used in this case. The $p<0.05$ level was used as the level of statistical significance. In the tables, p-values >0.20 are excluded to avoid information overload.

TABLE I. Number of vehicles in 1997 and 1999.

	1997		1999	
Main road (Västra Bräckevägen)				
No. of vehicles: Total (Heavy)				
Day and night	24 600	(4600)	2400	(300)
Day (06-22)	23 225	(4475)	2200	(270)
Night (22-06)	1375	(125)	180	(31)
Local street (Dysiksgatan)[a]				
Total number of vehicles day and night	10 100		4600	

[a]Local street between exposed and control area, speed limit 50 km/h.

IV. RESULTS

The main results of this study are presented below. A full report of the results is given in Report 4/99 (Öhrström and Skånberg, 1999).

A. Noise levels before and after reduction in road traffic

With the completion of the tunnel road traffic on the main road (Västra Bräckevägen) fell by 90% between 1997 and 1999. This road was also reconstructed: it was narrowed from 24 m to 12 m, and green areas and trees were planted on both sides. Table I shows the number and type of vehicles in 1997 and 1999. L_{Amax} levels were unchanged between 1997 and 1999 (87 dB). This means that the indoor noise levels [estimated from a facade reduction ($R'45°$,w) of 31 to 34 dB] from single noise events of 53–56 dB L_{Amax} occurred rather frequently during the daytime and to a lesser extent during the night time also in 1999 in the exposed area. $L_{90night}$ was 37 dBA and did not change between 1997 and 1999, whereas L_{90day} and L_{01day} decreased by 12 dBA from 52 and 72 dBA, respectively, in the exposed area.

TABLE II. Calculated outdoor and indoor traffic noise levels in dB $L_{Aeq,24h}$ in 1997 and 1999.

	Exposed area $L_{Aeq,24h}$ (min and max levels)		Control area $L_{Aeq,24h}$ (min and max levels)	
	1997	1999	1997	1999
Northwest facade (noisy side)	56–69	44–57	40–52	39–48
Southeast facade (quiet side)	48–64	38–50	40–51	40–48
Outside bedroom	43–67	39–55	40–51	39–48
Outside living room	48–67	38–55	40–51	39–48
Inside bedroom[a]	<20–36	<20–24	<20	<20
Inside living room[a]	<20–36	<20–24	<20	<20

[a]The value <20 is given if outside traffic noise level-a facade reduction of 31 or 34 dB = lower than 20 dB.

Table II shows calculated noise levels from road traffic noise in $L_{Aeq,24h}$ outdoors and estimated noise levels indoors. Levels are given as minimum and maximum values in each area. The table shows a decrease in outdoor $L_{Aeq,24h}$ levels of 10–14 dB in the exposed area in 1999 after the reduction in road traffic. With the exception of one house, all dwellings in the exposed area had an ''acceptable'' environment in terms of noise after traffic reduction, i.e., the noise level recommended by the Swedish Parliament ($L_{Aeq,24h}=55$ dB) was no longer exceeded. Calculated $L_{Aeq,24h}$ levels in the control area were unchanged or decreased by 1 dB (in one case—4 dB). Indoor noise levels in $L_{Aeq,24h}$ due to road traffic with closed windows were estimated to be low in the exposed area after the traffic reduction in 1999. Noise from single vehicles could still, however, be clearly heard indoors when the windows were closed.

Results of measured and calculated $L_{Aeq,24h}$ noise levels in dB showed fairly good agreement (−1 to 3 dB for measured levels) at the measurement sites near the main road in the exposed area. There were, however, large discrepancies between measured and calculated noise levels at measurement sites in the control area (+2 to 9 dB for measured levels). The Nordic Prediction Method gives accurate predictions on the directly exposed side. However, inside courtyards, and in areas with shielding buildings and at long distances this prediction method gives large errors (Kihlman *et al.*, 2002).

B. Perception and use of living environment

In 1997, before the changes were made in road traffic, a majority of residents in the exposed and the control areas were *satisfied with and liked their outdoor environment*, although significantly more ($p<0.002$) in the control area (86% very satisfied in the exposed area versus 92% in the control area). The residents' enjoyment of their outdoor environment in the exposed area improved significantly ($p=0.02$) after the traffic change. In 1997, a significantly ($p<0.0001$) higher proportion (40%) of the residents in the exposed area wanted to *move out of the area*, mostly for environmental reasons. There were no significant differences between the areas in 1999, however, and only 13% of the residents in the exposed area mentioned environmental reasons for wanting to move away from the area. When the residents were asked to compare their enjoyment before and after the building of the Lundby tunnel (''better,'' ''no difference,'' ''worse'') a significantly larger proportion in the exposed area felt they *enjoyed their neighborhood* more (71%, $p<0.001$) in 1999 than in 1997. This was also the case in the control area although the improvement was not as dramatic (21%).

Comparisons of effects by road traffic noise with respect to the use of the home environment

When the residents were asked if they used their homes and outdoor environment in a different way after the Lundby tunnel was opened in 1998, a significantly higher proportion (51%, $p<0.01$) of those living in the exposed area answered ''yes,'' as opposed to 27% in the control area. In the exposed area 40% of the respondents spent more time in their garden than before and 45% kept their windows open more often in comparison with the control area (1% and 5%, respectively). These differences between the exposed area and the control area were statistically significant, $p<0.001$.

The comments to the open question ''*Do you think that the way you use your house and the environment has been affected by the building of the tunnel*?'' revealed that the changes in road traffic affected the residents in several ways. Fifty-one percent in the exposed area and 27% in the control area answered yes to this question and also explained in what way. The comments given by the respondents in the exposed area mostly concerned the extensive reduction in exhaust fumes and noise and the fresh air. It had become easier to speak with one another outdoors because it had become quieter and calmer. It had become possible to enjoy being outdoors and to hear birds. The comments given in the control area dealt more with the environment as a whole, e.g., nicer, safer for children, better environmental quality.

TABLE III. Annoyance caused by road traffic noise in 1997 and 1999.

	Exposed area		Control area	
	1997	1999	1997	1999
$L_{Aeq,24h}$-levels (average calculated)	67	55	45	44
Number of respondents	50	45	92	75
Noise annoyance (%)				
Not at all annoyed	0	26.7	47.8	54.7
Not very annoyed	4	44.4	39.1	44
Rather annoyed	38	22.2	11.9	1.3
Very annoyed	58	6.7[b]	1.1	0
Noise annoyance scale 0–10 (mean)	8.9	1.4[b]	2.3	1.7[a]

[a]$p<0.01$.
[b]$p<0.001$.

C. Annoyance from road traffic and other environmental nuisances

The dominant sources of annoyance (expressed as percentage rather and very annoyed) in 1997 in the exposed area were related to road traffic: noise (96%), exhaust fumes (94%), and vibrations (82%); followed by dust (57%) and odours (48%) from industries or other sources. Annoyance caused by odors from industries was dominant in the control area (27%), followed by dust (14%), noise from road traffic (13%), and exhaust fumes (13%).

Road traffic noise was estimated to be more annoying during certain periods of the day and night by most of the respondents in the exposed area and by half of the respondents in the control area. Noise was most annoying between 6:00 and 8:00 am and between 12:00 pm and 6:00 pm, according to 86 and 80%, respectively, in the exposed area and 34 and 41%, respectively, in the control area. Eighteen percent in the exposed area and 3% in the control area found noise more annoying during the evening (18–22). Only 6% of the respondents in the exposed area and 1% in the control area found noise more annoying during the night (22–06).

Table III shows the results on annoyance caused by road traffic noise in 1997 and 1999. The table shows a highly significant decrease in the exposed area ($p<0.001$) both in the percentage of very annoyed (from 58% to 6.7%) respon-

TABLE IV. Activity disturbances due to road traffic noise in the exposed area in 1997 and 1999.

Percentage rather or very annoyed by road traffic noise during the following activities:	Exposed area		Control area	
	1997 $n=50$	1999 $n=45$	1997 $n=92$	1999 $n=75$
Indoor activities				
Conversation/telephone conversation	15	0	0	0
Radio/TV	12	2.2	1.1	0
Concentration	12	2.2	2.2	0
Rest/relaxation	38	11.1	5.4	0
Difficulties in falling asleep	24	8.9	2.2	0
Awakenings	28	8.8	3.3	0
Does not keep windows open as often as desired	58	4.4	3.3	0
Outdoor activities				
Able to be on the terrace or patio	38	4.4	1.1	0
Conversation	48	8.8	0	0
Relaxation	50	11.1	2.2	0

dents and on the numerical annoyance scale graded from 0 ''not at all annoyed'' to 10 ''extremely annoyed.'' The size of this extent of annoyance agrees well with expected noise annoyance at $L_{Aeq,24h}=55$ dB. A significant decrease in annoyance ($p<0.01$) was also seen on the numerical scale in the control area with a relatively unchanged noise level in $L_{Aeq,24h}$. The correlation in 1997 based on individual data in both areas ($n=142$) between noise annoyance and $L_{Aeq,24h}$ levels at the most exposed facade was highly significant ($r=0.76, p<0.0001$). In 1999 a very low correlation was found due to the small variation in individual data both for noise levels and for annoyance.

D. Activity disturbances by road traffic noise

In the control area, 5% at most reported that road traffic noise affected various activities in 1997, and none reported this in 1999. Results for the exposed and the control area are shown in Table IV. In 1997, ''not being able to keep windows open as often as one wished to'' was the most annoying effect of road traffic noise, followed by ''outdoor disturbances,'' ''relaxation,'' ''communication,'' and ''not being able to spend time on the terrace or patio.'' Activity disturbances indoors were dominated by ''disturbance of rest and relaxation,'' ''awakenings,'' and ''difficulties in falling asleep.'' In 1999 after the traffic noise reduction, all types of activity disturbances were significantly lower ($p<0.01$) in the exposed area.

The residents were also asked to compare the effects of road traffic noise on activities with the effects one year earlier before the Lundby tunnel was opened (''less,'' ''no difference,'' and ''more''). Between 60% and 84% in the exposed area reported less extensive effects caused by road traffic noise. The greatest decrease was in outdoor activities, followed by conversation, relaxation, and listening to radio/TV indoors. All changes were significant ($p<0.0001$).

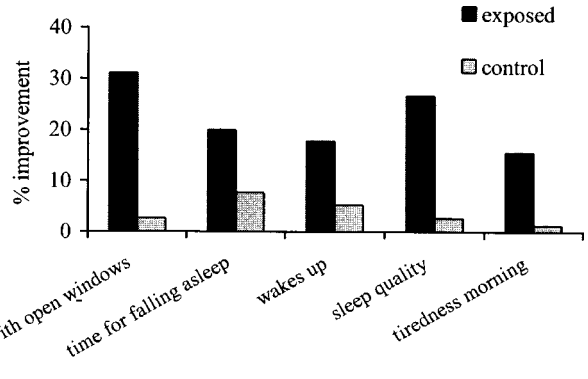


FIG. 2. Reported improvement in sleep after changes in road traffic.

E. Sleep disturbances by road traffic noise

The residents were asked to compare their sleep and sleep behavior in 1999 with their sleep one year earlier before the Lundby tunnel was opened (see results in Fig. 2). The only significant change in the control area was that 8% of the residents reported that it had become easier to fall asleep. A significant proportion of the respondents in the exposed area estimated that their sleep had changed for the better with: ''windows open more often during the night'' (31%), ''better sleep quality'' (27%), ''easier to fall asleep'' (20%), ''fewer awakenings'' (18%) and ''less tired in the morning'' (16%). All changes were statistically significant (waking up less frequently $p=0.02$ and other parameters $p<0.01$).

In answer to the direct, open question ''*Do you think that your sleeping habits or your sleep have been affected by the building of the Lundby tunnel*,'' 24% in the exposed area versus 3% in the control area answered positively ($p<0.001$). The reasons for the change in sleep and sleeping habits reported by these respondents in the exposed area were: ''it is quieter now;'' ''there is less traffic to disturb sleep;'' ''it feels like a move to the country;'' ''it is possible to sleep without the water bed vibrating with the traffic;'' ''it is better when there are no trucks to wake one up;'' ''calmer mornings;'' ''not waking up because of trucks driving in the morning between five and six o'clock;'' ''it is possible to sleep longer in the morning without being woken up;'' and ''being sick less often.''

F. Psycho-social well-being

The prevalence of different physiological and psychological symptoms was assessed as an indicator of psycho-social well-being. Figure 3 shows the prevalence of symptoms in the exposed area and in the control area reported a few times a week or every day in 1997, before the road traffic changes. All symptoms, except headaches, were more common in the exposed area than in the control area before the changes in road traffic. Significant differences between the two areas were present for ''unsociable/wants to be alone'' ($p=0.03$), ''uncomfortable/upset stomach'' ($p=0.02$) and ''unhappy and depressed'' ($p=0.02$). The total score for psycho-social well-being was significantly better in the control area than in the exposed area, mean=12.1 versus mean=13.9, $p=0.007$. Table V shows the prevalence of

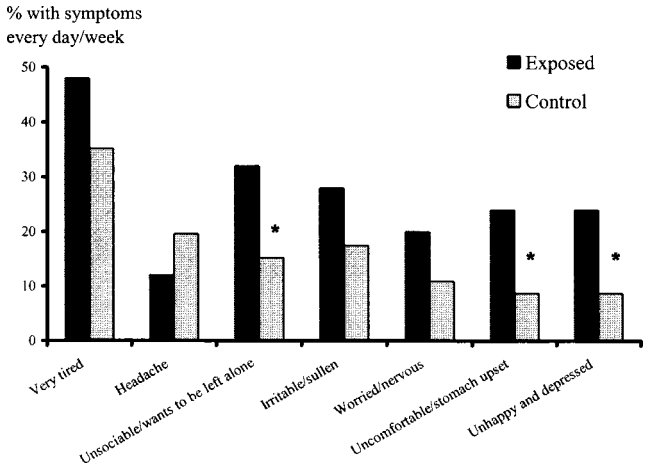


FIG. 3. Prevalence of symptoms in the exposed and the control areas in 1997.

symptoms reported a few times a week or every day in both areas in 1997 and 1999. Several symptoms decreased significantly in the exposed area (Wilcoxon one-sided test) after the change in road traffic (uncomfortable/upset stomach $p=0.04$, unsociable/wants to be left alone $p=0.048$, irritable/sullen $p=0.02$ and worried/nervous $p=0.03$). In 1999 the differences in prevalence of symptoms between the two areas were reduced for six out of the seven symptoms and were no longer significant. A significant, not hypothesized, decrease in symptoms was also seen in the control area for ''unhappy and depressed'' and ''worried/nervous'' ($p=0.02$ and 0.04, respectively, Wilcoxon two-sided test).

The respondents' own comparisons of how often they experienced the various symptoms as compared with 1998 before the Lundby Tunnel was opened (more often, no difference, less often) showed no significant changes for any of the symptoms. However, in response to the open question as to whether the resident believed that the tunnel had affected general well-being in any way, 50% of the respondents in the exposed area said ''yes'' as compared with 25% in the control area ($p=0.01$). The motivations given by residents of the *exposed area* for the tunnel's effect on general well-being were several: ''it is quieter and calmer and gives more peace in the garden;'' ''we enjoy the garden more;'' ''it is calmer, more enjoyable;'' ''it is quieter;'' ''something good has happened in the neighborhood;'' ''the general environment has improved;'' ''traffic noise and emissions have decreased;'' ''there is better air, little traffic, a more pleasant street;'' ''it was never pleasant to sit in the garden before;'' ''the environment is better, there is less noise, emissions and vibrations;'' ''less dirt in the air;'' ''the air feels cleaner and doesn't smell like exhaust fumes unlike before;'' ''no disturbance by noise and exhaust fumes now;'' ''no more irritation over queues of traffic and cars idling;'' ''sleep better, good sleep is good for everything and everybody;'' ''more relaxed at home, freedom to be outside, happy;'' ''feel better because of the cleaner air, happy to hear birds;'' ''less noise and exhaust fumes, makes you feel better;'' ''yes (greater well-being) in the long run.''

The motivations given by the residents of the *control area* for the effect of the Lundby tunnel on general well-being were: ''the standard of the entire area has been raised;'' ''the area feels more connected;'' ''cleaner air and quieter;'' ''calmer and quieter outside;'' ''it feels good that it was built;'' ''knowing that traffic has been moved farther away is positive;'' ''better environment and less traffic;'' ''knowing that all traffic goes through the tunnel,'' ''the area has become prettier;'' ''better environment;'' ''less noise and vibration;'' ''nice that the heavy traffic is gone;'' ''... (well-being) has not been affected a great deal but there has been a positive development in the total environment;'' ''I feel better in my home;'' ''fresher air;'' ''more the countryside;'' ''yes (well-being has been affected) in the long run;'' ''my tinnitus has become better with less car honking;'' ''had migraines during the construction with dry noise and throat, better now.''

Relationship between psycho-social well-being, annoyance, sensitivity to noise, and other individual factors

Table VI shows the correlation between the seven psychological and physiological symptoms, annoyance, and sensitivity to noise. Annoyance caused by road traffic noise was significantly related to all the symptoms ($r_s=0.24$ to 0.28) except headache ($r_s=-0.02$). Sensitivity to noise was significantly ($r_s=0.28$ to 0.37) related to all symptoms. The absence of a correlation between headaches and noise annoyance may explain why headaches was the only symptom that

TABLE V. Prevalence of physiological and psychological symptoms before and after changes in road traffic.

	Exposed area			Control area		
Symptoms reported a few times per week or every day (%)	1997 $n=50$	1999 $n=45$	97/99 p-value	1997 $n=92$	1999 $n=75$	97/99 p-value
Headaches	12	13.3	···	19.6	17.3	···
Very tired	48	31.1	0.09	35.8	29.4	0.20
Uncomfortable/upset stomach	24	15.6	0.04	8.7	8	···
Unhappy and depressed	24	11.1	0.20	8.7	5.3	0.02
Unsociable/wants to be alone	32	20	0.048	15.2	12	···
Irritable/sullen	28	17.8	0.02	17.4	10.7	0.08
Worried/nervous	20	15.6	0.03	10.9	5.3	0.04
Total psycho-social well-being (mean value)[a]	13.9	12.4	0.03	12.1	11.5	0.08

[a]=sum of scores (1–4) on all symptoms. Higher value represents worse psycho-social well-being.

TABLE VI. Relationship (Spearman r_s) between well-being, annoyance from road traffic noise, and noise sensitivity (results from 1997, $n=141$).

Psychological and physiological symptoms	Annoyance from road traffic noise r_s	Sensitivity to noise r_s
Headaches	−0.02	0.26[b]
Very tired	0.26[b]	0.35[b]
Uncomfortable/upset stomach	0.24[a]	0.37[b]
Unhappy and depressed	0.24[a]	0.29[b]
Unsociable/wants to be left alone	0.26[b]	0.28[b]
Irritable/sullen	0.24[a]	0.30[b]
Worried/nervous	0.28[b]	0.31[b]

[a]*P*-value: $p<0.01$.
[b]*P*-value: $p<0.001$.

was not more frequent in the exposed area than in the control area. The different symptoms were closely interrelated, although headaches had the lowest correlation with the other symptoms. Some significant relationships were found between various symptoms and age and long-term disease. The higher the age the *less* frequent were the symptoms of ''irritable/sullen'' ($r_s=-0.28$) and ''unsociable/wants to be left alone'' ($r_s=-0.22$). The prevalence of long-term disease varied significantly with ''headache'' and with feeling ''very tired'' ($r_s=0.21$ and 0.24, respectively) but was not related to the other symptoms.

V. DISCUSSION

A. Comparisons with the previous studies in the residential areas in 1986 and 1987

The overall results on adverse effects of noise obtained in 1997 in this longitudinal study, before the extensive changes in road traffic were executed, support the findings of the earlier studies in these residential areas performed in 1986 and 1987 (Öhrström, 1989; Björkman *et al.*, 1988). Sixty-three percent of respondents were very annoyed by road traffic noise in 1986 and the extent of annoyance was still high, 58%, in 1997. The frequency of noise-disturbed activities (e.g., conversation, relaxation, and sleep) varied between 20% and 51% in 1986 and was only marginally lower, 12% to 38%, in 1997. The prevalence of psychological and physiological symptoms showed a similar pattern over the years from 1986 to 1997. The symptoms were, on average, three times more common in the exposed area than in the control area in 1986, and twice as common in the exposed area in 1997. The frequency of symptoms varied between 20% (worried/nervous) and 68% (very tired) in the exposed area in 1986 and between 12% (headaches) and 48% (very tired) in 1997. The slight decrease in adverse effects and in symptoms between 1986 and 1997 might have to do with noise abatement measures (ban on heavy vehicles during the night, noise-reducing asphalt and a lower speed limit) undertaken in 1987. On the other hand, there were no indications of a reduction in adverse effects or in noise levels in the study undertaken five months after these measures in 1987 (Björkman *et al.*, 1988). It should be noted, however, that the results are not completely comparable over this 10-year period since no data on personal identification are allowed to be kept for so long. Also, a small number of respondents in apartment buildings were included in the exposed area in the 1986 study. Nevertheless, a fair interpretation of these results is that there is little habituation, if any, for people who have been exposed to a high amount of pollution from road traffic for a long period of years.

B. General annoyance and indoor and outdoor activity disturbances

An outdoor noise level of 55 dB $L_{Aeq,16h}$ is recommended by WHO (WHO, 2000) to protect from serious annoyance and 50 dB $L_{Aeq,16h}$ to protect from moderate annoyance. Outdoor noise levels in this study were ca. 67 dB $L_{Aeq,24h}$ (i.e., 69 dB $L_{Aeq,16h}$) at the most exposed facade and ca. 62 dB $L_{Aeq,24h}$ (i.e. 64 dB $L_{Aeq,16h}$) at the quieter facade. The extent of annoyance (58% very annoyed) in this study is thus much higher than in the dose-response curve for road traffic noise derived by Miedema and Vos (1998), although annoyance at the lower levels agrees well with their curve. The higher extent of annoyance at higher noise levels might be due to some extent to the additional exposure from vibrations induced by heavy vehicles (long-distance trucks) and attenuated by the soft ground (deep clay). Annoyance caused by vibrations was widespread: 82% were rather or very annoyed by vibrations from road traffic. This is in line with results from socio-acoustic surveys on noise and vibrations from trains (Öhrström and Skånberg, 1996) where it was found that annoyance caused by noise increases strongly in the presence of vibrations induced by the trains. It should also be noted that annoyance to exhaust fumes was very high: 94% were rather or very annoyed by this. Fields' (1993) findings also support the hypothesis that an awareness of non-noise problems increases noise annoyance.

General annoyance is sometimes regarded as a vague and not very serious noise-induced adverse health effect, whereas disturbance of activities (e.g., on speech interference, sleep, and relaxation) are clearly understandable effects and considered as more ''hard data.'' The relatively lower occurrence of disturbances of activities by noise in this study in comparison with the high general annoyance (96% rather or very annoyed) caused by noise might also have to do with attitudes. For example, it is a well known fact that noise from road traffic is more annoying than noise from trains although train noise causes greater speech interference than road traffic noise (Moehler, 1988; Öhrström and Skånberg, 1996; Miedema and Vos, 1999). One explanation offered is a more positive attitude towards trains than to the less environmental friendly road traffic.

The results in this study suggest that the perception of noise and the extent of its disturbing effects *outdoors* contribute more to the general annoyance reaction than indoor activity disturbances. Also, not being able to keep windows open as often as one wished to caused widespread (58%) disturbance. Outdoor activities were more frequently reported to be disturbed by road traffic, e.g., 60% did not even like to be outdoors on their patio due to the road traffic noise and, if they did, most of them felt rather or very annoyed. The large majority, 80%, also reported that they were affected by noise during conversation and relaxation and the

majority of them were rather or very annoyed (ca. 50% of all respondents) by noise during these activities outdoors.

The relatively lower occurrence of indoor activity disturbances, e.g., 16% for conversation and 38% for relaxation may partly be because windows were mostly kept closed. Furthermore, a large percentage of the residents in the exposed area had insulated their facades and/or windows. Another situational factor that might explain the relatively moderate occurrence of reported disturbances of sleep, such as difficulties falling asleep and awakenings (ca. 25%) and relaxation (38%) due to noise, is that a majority, 84%, had bedroom windows facing the quieter facade of the dwelling. Still, the noise levels indoors were in all cases higher than levels recommended for protection against sleep disturbances by WHO (WHO, 2000) (indoors 45 dB L_{Amax} and 30 dB $L_{Aeq,8h}$ and, for protection against sleep disturbances with open windows, L_{Amax}60 dB and L_{Aeq8h}45 dB outdoors). In this study indoor noise levels for L_{Amax} (closed window) in bedrooms and living rooms facing the most exposed facade were ca. 55 dB, and ca. 50 dB if these rooms were facing the quieter facade. With open windows the noise level would increase by 10 dB to 65 and 60 dB L_{Amax}, respectively. The importance of situational factors for annoyance reactions was pointed out by Fields (1993), who found that insulation has a substantial effect on annoyance from noise. Other studies have shown that sleep disturbances due to road traffic noise are reported more frequently when bedroom windows face the most noisy facade (Langdon and Buller, 1977; Öhrström, 1993; Lercher, 1996). A healthy sound environment in residential areas where few people are annoyed or disturbed by noise during various activities is probably achieved only where there is also access to quiet places outdoors.

C. General well-being and psycho-social symptoms

As a consequence of noise-induced annoyance and disturbed activities including sleep, general well-being may be substantially reduced. Support for a causal relationship between noise exposure and psycho-social well-being is, however, limited (Health Council of the Netherlands, 1994) and, as Lercher stated (1996) in his extensive paper on the potentials and limitations of major conceptual research models, ''firm conclusions about long-term consequences for health are still hampered by inconsistent results.'' In this study psycho-social well-being was evaluated by the same variables that were used together with a questionnaire on mood in the study in 1986 (Öhrström, 1989). The variables were chosen to reflect symptoms of stress, dissatisfaction, and sleep deprivation and were highly correlated with variables in the mood adjective checklist (Sjöberg *et al.*, 1977) used in 1986. These types of effects are often subtle and indirect and many of the psycho-social effects are assumed to result from the interaction of a number of non-auditory variables (WHO, 2000). To reduce the number of questions in the questionnaire and, as knowledge and methods in this field are limited, open questions were used as an attempt to gather information of how the respondents experienced their living environment and how they generally felt. The problem with these types of questions is that they are difficult to handle statistically. Still these questions give important, additional new knowledge.

The results of the present study on the prevalence of psychological and physiological symptoms in 1997 support the previous findings from 1986 (Öhrström, 1989). The total score on psycho-social well-being, i.e., the sum of the seven symptoms, was significantly poorer ($p<0.01$) in the exposed area than in the control area. All symptoms except ''headaches'' were reported more frequently in the exposed area than in the control area in 1997 and three of these (''uncomfortable/upset stomach,'' ''unsociable/wants to be left alone,'' and ''unhappy and depressed'') reached statistical significance whereas the symptoms ''irritable/sullen,'' ''worried/nervous,'' and ''very tired'' did not. Headaches, the only symptom that was more frequent in the control area, was, however, in contrast to the other symptoms, not correlated with noise annoyance and was probably not linked to exposure to noise. These types of effects are often subtle and indirect and many of the psycho-social effects are assumed to result from the interaction of a number of nonauditory variables (WHO, 2000). In this study both ''headache'' and ''feeling very tired'' were significantly related to long-term disease. This was, however, not the case for the other symptoms.

D. Effects of changes in road traffic exposure

The dramatic reduction in road traffic on the main road from ca. 25 000 vehicles to 2400 vehicles per 24 hours, of which ca. 20% were heavy vehicles and long-distance trucks, led to a decrease in noise levels from 67 to 55 dB $L_{Aeq,24h}$. The L_{Amax} levels remained high, 87 dB. Also, the number of vehicles on the local road between the exposed and the control area decreased from ca. 10 100 to 4600. The reduction in road traffic led to a large decrease in annoyance from 51% very annoyed at 67 dB to 7% very annoyed at 55 dB, which is what can be expected in stable situations according to other studies, (e.g., Miedema and Vos, 1998, 1999). The respondents already knew in 1996 about the plans to reduce traffic and had thus no reason to exaggerate their annoyance in 1997. The large reduction in noise annoyance can thus not be interpreted as an overreaction to a change in noise exposure as has sometimes been reported in the literature (e.g., Fields, 1993). The large decrease in *annoyance* after traffic noise reduction may, however, not solely be explained by a decrease in noise levels. Annoyance from exhaust fumes decreased in a similar way as annoyance from noise and the same was true for annoyance from vibrations. A contributing factor behind the extensive reduction in noise annoyance may also be associated with changes in the appearance of the area and the improved overall satisfaction with the residential area. When the new traffic tunnel was built the old main road was reconstructed from a 24-m-wide highway to a 12-m-wide small local road. Trees were planted on both sides, which resulted in a much more attractive appearance and entrance to the area. This would be in agreement with Langdon (1976), who concluded that the appearance of the residential area was important for the perception of dissatisfaction with road traffic noise. A somewhat unexpected finding is the decrease in annoyance from 13% to 1% rather or very

annoyed in the control area. There are, however, several possible explanations for this change. Even if the calculated L_{Aeq} was only slightly reduced, the number of vehicles on the nearby local road was reduced. The residents in the control area had to pass through the main road both before and after it was rebuilt to a small road to get to their homes and were thus very aware of the changes in road traffic. Furthermore, the large discrepancy between measured and calculated levels indicate that the noise levels in the control area were somewhat higher than calculated. As the measurements were not supervised, noise from other sources could have contributed to the noise levels. It is also known (Kihlman *et al.*, 2002) that other more distant roads as well as reflections of noise from nearby buildings (not considered in the Nordic Prediction Method) contribute to the measured noise levels. The fact that 21% of the respondents in the control area enjoyed their neighborhood more after the changes in road traffic and found it safer and more attractive probably also contributed to the decrease in annoyance. Fields' (1993) conclusions from the eight studies he reviewed that a reduction in noise levels, even at relatively low noise levels below 55 dB DNL, may lead to a decrease in annoyance are also in line with this.

The large decrease in road traffic and in noise-induced annoyance was associated with a significant reduction of the extent of *activity disturbances* both outdoors and indoors. A large improvement was seen for outdoor activities and also for indoor activities related to restoration and sleep as well as for a change in the habits of keeping windows open. These large improvements agree well with the change in disturbance demonstrated by the respondents' own comparison questions on activity disturbances, sleep disturbances, and tiredness during the morning. Before the changes in road traffic, early morning hours were mentioned by a large majority as the time of the day when road traffic caused the greatest annoyance. As the number of noise events during the night was reduced from ca. 1375 to 180 and the number of heavy vehicles (long-distance trucks) from ca. 125 to 30 per night, a large improvement in sleep was to be expected for those who had bedroom windows facing the main road or were previously hindered from sleeping with open windows (Öhrström, 2000). Results of cross-sectional studies on noise and sleep disturbances seldom show significant dose-effect relationships (e.g., Pearsons, 1995) and discrepancies between results obtained in laboratory studies and field studies are large for some sleep disturbance effects (e.g., awakenings) (Pearsons, 1995), whereas this is not the case for other sleep parameters (Langdon and Buller, 1977; Öhrström, 2000). Langdon and Buller (1977) found a significant relationship at the group level between $L_{10,22\text{-}06}$ and difficulties in falling asleep ($r=0.91$) and waking during the night ($r=0.84$) because of noise. In the present longitudinal study, the same individuals were studied before and after a reduction in road traffic noise, which makes the results more reliable than results obtained in cross-sectional studies.

The results on *general well-being* in terms of psychological and physiological symptoms after the changes in road traffic were less clear than the effects on annoyance and activity disturbances. After the reduction in road traffic a decrease in the prevalence of symptoms was seen in the exposed area for six of the seven symptoms. The total score on psycho-social well-being was significantly improved and this was also the case for the four symptoms ''uncomfortable/upset stomach,'' ''unsociable/wants to be left alone,'' ''irritable/sullen,'' and ''worried/nervous.'' The decrease in the symptoms that differed significantly between the exposed and the control areas in 1997—''unhappy and depressed'' and ''very tired'' was not statistically significant. In the control area a slight decrease in symptoms was seen, significant for ''unhappy and depressed'' and ''worried/nervous.'' There was some contradiction between the prevalence of symptoms and the results obtained from the respondents' own comparisons of how often they experienced the various symptoms as compared with 1998 before the Lundby Tunnel was opened (more often, no difference, less often). The response to these questions showed no significant changes for any of the symptoms. On the other hand, half of the respondents in the exposed area believed that their general well-being had been positively affected after the building of the tunnel. The most apparent interpretation of this is that the symptoms only measure a limited aspect of general well-being and that it is hard to remember how often one felt the symptoms one year ago. The comments from the residents on how they experienced their environment after the changes in road traffic and the environment indicates, however, that their general well-being was affected in a profound and positive way.

The positive outcome of the reduction in road traffic exposure, and of the large change in appearance of the areas, can not only be described in terms of a substantial reduction in annoyance, activity disturbances, and psycho-social symptoms. Also, more diffuse, but not necessarily less important, health implications such as better life quality (Taylor, 1998), e.g., a greater possibility to use and enjoy the living environment, were found. Noise from road traffic is associated with negative experiences and reactions such as annoyance, irritation, speech interference, and disturbed sleep. The absence of noise and combined effects with simultaneous exposures (e.g., exhaust fumes, vibrations, fear of accidents, appearance of the area) in the living environment is associated with positive reactions such as the ability to perceive sounds such as birds singing, being able to talk with neighbors and ''listening'' to the silence. It is clear from many of the comments made that the environment was perceived as more healthy and restorative and that the possibility to stay outdoors in their own garden was seen as an important part of life quality. The importance of personal (e.g., noise sensitivity), attitudinal and, at least to some extent, situational factors for effects of noise is relatively well established (e.g., Lercher, 1996) while new knowledge needs to be assessed concerning ''softer'' health effects associated with noise and other exposures caused by road traffic. These effects constitute clear indicators of improved health in the population after a reduction in exposure to road traffic. New methods should be developed for assessment of these types of health effects, e.g., based on qualitative analyses from in-depth interviews.

The long-term goal (year 2020) in Sweden is for noise levels to be below 55 dB $L_{Aeq,24h}$. A ''good and healthy'' sound environment is probably achieved only if noise levels

are much lower, around 40–45 dB $L_{Aeq,24h}$, levels that are seldom found in today's densely populated areas. According to this definition, the control area in this study could be characterized as a satisfactory and healthy sound environment. Very extensive noise abatement, like the building of the Lundby tunnel, is seldom achievable. One of the questions currently being studied in a large, multi-disciplinary Swedish research program (Berglund *et al.*, 1999) is whether a healthy sound environment can be achieved if one side of the dwelling has high noise levels above 60 dB $L_{Aeq,24h}$, while at least one other side of the dwelling has low noise levels around 40–45 dB $L_{Aeq,24h}$.

VI. CONCLUSIONS

It is concluded from these long-term investigations before and after extensive changes in road traffic that exposure to a high amount of road traffic induces adverse effects, not only in terms of annoyance and disturbance of daily activities indoors and outdoors, but also in terms of lower psychosocial well-being. It is suggested that both noise and other pollutants from road traffic contribute to these effects.

To be able to use the outdoor environment and to have the possibility to keep windows open is essential for general well-being and daily behavior and this implies that access to both a quiet indoor and outdoor section of the residency is of importance for achievement of a healthy sound environment.

For the creation of healthier sound environments, the simultaneous knowledge of long-term health consequences of exposure to noise and pollutants from road traffic and further studies are needed. These studies could, preferably, be performed in connection with traffic mitigation and should focus more on ''softer'' health outcomes and well-being than hitherto and not only on ''traditional'' medical outcomes.

ACKNOWLEDGMENTS

The Swedish National Road Administration and the Swedish Regional Road Administration financially supported this research. Ingemansson Technology AB, Göteborg, performed calculations of noise levels and measured facade reduction. Martin Björkman, Göteborg University, performed other noise measurements. Annbritt Skånberg, Agneta Agge, and Gunilla Arvidsson from Göteborg University assisted in the field work.

Berglund, B., Kihlman, T., Kropp, W., and Öhrström, E. (**1999**). ''Soundscape support to health,'' Stockholm: *A MISTRA Research Programme*, 1999 http://www.soundscape.nu/.

Björkman, M., Levein, B., Rylander, R., and Öhrström, E. (**1988**). ''Effects of road traffic noise for the population along Bräckeleden in Gothenburg,'' Report 8/88, Department of Environmental Medicine, Göteborg University, ISSN-0280-2600 (in Swedish).

Fields, J. (**1993**). ''Effect of personal and situational variables on noise annoyance in residential areas,'' J. Acoust. Soc. Am. **93**, 2753–2763.

Health Council of the Netherlands (**1994**). ''Noise and Health. The Hague: Health Council of the Netherlands,'' Report by a Committee on Noise and Health.

IEH (**1997**). ''The Non-Auditory Effects of Noise,'' Leicester, UK, Institute of Environment and Health, IEH Report R10.

ISO (**1998**). ISO 140-5 Acoustics—Measurement of sound insulation in buildings and of building elements—Part 5: Field measurements of airborne sound insulation of façade elements and façades.

Kihlman, T., Öhrström, E., and Skånberg, A. (**2002**). ''Adverse health effects and the value of access to quietness in residential areas,'' in *Inter-Noise 2002*, edited by A. Selamet, R. Singh and G. C. Maling, Dearborn, MI, International Institute of Noise Control Engineering (I-INCE), Paper N 484 (available on CD).

Langdon, F. J. (**1976**). ''Noise nuisance caused by road traffic in residential areas: Part I,'' J. Sound Vib. **47**, 243–263.

Langdon, F. J., and Buller, I. B. (**1977**). ''Road traffic noise and disturbance to sleep,'' J. Sound Vib. **50**, 13–28.

Lercher, P. (**1996**). ''Environmental noise and health: an integrated research perspective,'' Environ. Int. **22**, 117–129.

Miedema, H. M. E., and Vos, H. (**1998**). ''Exposure-response relationships for transportation noise,'' J. Acoust. Soc. Am. **104**, 3432–3445.

Miedema, H. M. E., and Vos, H. (**1999**). ''Factors that modify annoyance,'' J. Acoust. Soc. Am. **105**, 3336–3344.

Moehler, U. (**1988**). ''Community response to railway noise: A review of social surveys,'' J. Sound Vib. **120**, 321–332.

Öhrström, E. (**1989**). ''Sleep disturbance, psycho-social and medical symptoms—a pilot survey among persons exposed to high levels of road traffic noise,'' J. Sound Vib. **133**, 117–128.

Öhrström, E. (**1993**). ''Long-term effects in terms of psycho-social well-being, annoyance and sleep disturbance in areas exposed to high levels of road traffic noise,'' in *Noise and Man '93*, Noise as a Public Health Problem, Proceedings of the 6th International Congress, edited by M. Vallet, Institut National de Recherche sur les Transports et leur Sécurité Nice, France 5–9 July 1993, Vol. 2, pp. 209–212.

Öhrström, E. (**2000**). ''Sleep disturbances caused by road traffic noise—studies in laboratories and field,'' Noise Health **8**, 71–78.

Öhrström, E., and Björkman, M. (**1988**). ''Effects of noise-disturbed sleep—a laboratory study on habituation and subjective noise sensitivity,'' J. Sound Vib. **122**, 277–290.

Öhrström, E., and Skånberg, A. (**1996**). ''A field survey on effects of exposure to noise and vibrations from railway traffic. Part I: Annoyance and activity disturbance effects,'' J. Sound Vib. **193**, 39–47.

Öhrström, E., and Skånberg, A. (**1999**). ''Consequences of the Lundby Tunnel. Part I: Investigations of the experience of disturbance, sleep, health and well-being among the population living near Västra Bräckevägen, Göteborg. Part II: Detailed sleep investigations,'' Report 4/99, Department of Environmental Medicine, Göteborg University (in Swedish).

Passchier-Vermeer, W. (Ed.) (**1996**). ''Effects of noise on health,'' in *Noise/News International*, 1996 (September), pp. 137–150 (Chapt. 3 of a report on Noise and Health prepared by a Committee of the Health Council of The Netherlands).

Pearsons, K. S., Barber, D. S., Tabachnick, B. G., and Fidell, S. (**1995**). ''Predicting noise-induced sleep disturbance,'' J. Acoust. Soc. Am. **97**, 331–338.

Road traffic noise: Nordic prediction method (**1996**). Jonasson, H., and Lyse Nielsen, H. Copenhagen, Nordiska Ministerrådet, TemaNord 1996:525, ISBN/ISSN 92-9120-836-1.

Sjöberg, L., Svensson, E., and Persson, L. O. (**1977**). ''The measurement of mood,'' Scand. J. Psychol. **19**, 1–18.

Swedish Standard SS 02 52 54. (**1980**). Measurement of sound insulation in buildings and of building elements. International standards valid for Swedish standards.

SS EN ISO 717-1. (**1996**). Building acoustics—Evaluation of sound insulation in buildings and of buildings elements—Part I Airborne sound insulation.

Technical Specification ISO/TS15666 (**2003**). First edition 2003-02-01 Acoustics—Assessment of noise annoyance by means of social and socio-acoustic surveys. Reference number ISO/TS 15666:2003(E).

Taylor, S. M. (**1998**). ''Noise as a population health problem,'' in *Inter-Noise 98*, edited by V. C. Goodwin and D. C. Stevenson, (NZ Acoustical Society, Christchurch, New Zealand), Vol. 2, pp. 1157–1162.

WHO. (**2000**). ''Guidelines for community noise,'' edited by B. Berglund, T. Lindvall, D. Schwela and K. T. Goh (WHO, 2000).

External and internal noise surveys of London primary schools

Bridget Shield[a)]
Faculty of Engineering Science and Technology, London South Bank University, 103 Borough Road, London SE1 0AA, United Kingdom

Julie E. Dockrell
Psychology and Human Development, Institute of Education, London University, 25 Woburn Square, London WC1H 0AA, United Kingdom

(Received 13 December 2002; revised 17 October 2003; accepted 3 November 2003)

Internal and external noise surveys have been carried out around schools in London, UK, to provide information on typical levels and sources to which children are exposed while at school. Noise levels were measured outside 142 schools, in areas away from flight paths into major airports. Here 86% of the schools surveyed were exposed to noise from road traffic, the average external noise level outside a school being 57 dB L_{Aeq}. Detailed internal noise surveys have been carried out in 140 classrooms in 16 schools, together with classroom observations. It was found that noise levels inside classrooms depend upon the activities in which the children are engaged, with a difference of 20 dB L_{Aeq} between the ''quietest'' and ''noisiest'' activities. The average background noise level in classrooms exceeds the level recommended in current standards. The number of children in the classroom was found to affect noise levels. External noise influenced internal noise levels only when children were engaged in the quietest classroom activities. The effects of the age of the school buildings and types of window upon internal noise were examined but results were inconclusive. [DOI: 10.1121/1.1635837]

PACS numbers: 43.50.Qp, 43.50.Rq [DKW] Pages: 730–738

I. INTRODUCTION

There are several national and international guidelines relating to the acoustics of classrooms.[1–4] These mainly take the form of recommended values for reverberation time and background noise levels in teaching spaces, together with sound insulation requirements for schools. For example, the World Health Organization (WHO) Guidelines for Community Noise[2] specify an appropriate background level for classrooms as 35 dB L_{Aeq} during teaching sessions. This is also the level recommended in the USA in a recently approved ANSI standard.[3] In the UK legislation governing noise and acoustic conditions in schools was introduced in July 2003. The required acoustic standards are contained in Building Bulletin 93,[4] which specifies a maximum ambient level of 35 dB $L_{Aeq,30\,min}$ due to noise from sources such as ventilation plant, and intrusive external noise in unoccupied teaching spaces. Some standards also include guideline values for noise levels outside schools, for example, the WHO recommends that noise levels in school playgrounds should not exceed 55 dB L_{Aeq}, whereas in the UK Building Bulletin 93 specifies an upper limit of 60 dB $L_{Aeq,30\,min}$ at the site boundary and 55 dB L_{Aeq} in outdoor areas such as playgrounds and playing fields.

Despite the existence of guidelines for school and classroom noise, and a body of research on the effects of noise on children and teachers in the classroom, there is relatively little information on noise levels in classrooms and outside schools. The purpose of the study described here was to provide objective data on the external and internal noise environment of typical urban schools; and to investigate to what extent the external noise climate influences the noise levels inside schools and classrooms. Noise levels were measured outside 142 schools in London, England, and inside 16 schools, in approximately 200 locations, including empty and occupied classrooms, assembly halls, and corridors. In parallel with the noise surveys described, questionnaire surveys of children and teachers were carried out in over 50 schools, and the questionnaire results compared with the measured noise levels.[5]

II. BACKGROUND TO THE STUDY

Although there have been many previous studies concerning the effects of chronic noise exposure on children at school, there is relatively little published data quantifying the noise environment inside and outside schools. Previous work has shown that noise can have a detrimental effect upon the cognitive development of primary school children, and that older children in this group appear to be more affected than younger children.[6,7] Many of these studies have concluded that the chronic noise exposure of young children has a particularly detrimental effect upon their reading ability.[8,9] In general, it has been shown that aircraft noise has more effects than noise from other sources,[10–15] although effects on children's reading, attention, and long-term recall have also been found due to school exposure to train and road traffic noise.[16–21] Noise from road traffic has also been found to cause dissatisfaction with the classroom environment among teachers.[18] In a comparison of the effect of noise from different sources, it was found that aircraft and road traffic noise played at 66 dB(A) in the classroom affected long-term recall, whereas train noise had no effect.[22] Studies that have

[a)]Corresponding author. Electronic mail: shieldbm@lsbu.ac.uk

considered the effects of internal classroom noise[8,23–25] have found a significant drop in children's performance, particularly in learning to read, when the background noise level interferes with speech. By corollary, the reduction of background noise through acoustic treatment has been shown to improve the performance of both preschool[24] and primary school children.[25] Thus, overall, the evidence provided by previous studies shows that noise from a variety of sources inside or outside a primary school has a detrimental effect upon children's learning and performance at school, particularly reading, and that the effect can be greater with older children in this age range.[17,26]

Despite the proven significance of noise exposure at school, the number of schools affected by noise from different sources is unknown and there is relatively little data available on typical noise levels inside and outside schools. Examples of external noise levels outside schools in a densely populated urban environment are provided by surveys carried out in central Istanbul in the late 1990s,[27–29] which found levels ranging from 54 to 79 dB $L_{Aeq,5\,min}$,[27,28] and, in the most densely populated areas from 72 to 97 dB L_{Amax}; from 55 to 73 dB L_{A10}; and from 49 to 61 dB L_{A90}, all schools being subject to road traffic noise.[29]

With regard to noise levels inside schools, a further problem arises in interpreting previously published data, owing to the lack of a standard method for measuring noise in schools, and the difficulty of deciding what measurement represents a ''typical'' classroom noise level. These problems are acknowledged by Hodgson *et al.*[30] and by Picard and Bradley[31] in reviews, published in 1999 and 2001, respectively, of classroom noise surveys. There is a wide range of levels in the published data. For example, Hodgson *et al.*[30] in summarizing previous classroom noise surveys, dating from 1977–1991, found that classroom speech (that is, teacher) levels ranged from 40 to 80 dB(A); student activity levels from 40 to 70 dB(A); and ventilation noise levels in classrooms from 23 to 55 dB(A). Similarly, Picard and Bradley[31] noted that reported occupied levels in a full range of classrooms from kindergarten to university varied from 42 to 94 dB(A).

In a survey of university classrooms in Canada, Hodgson[32] found that a typical background noise level of 35 dB(A) in an empty classroom increased to 56 dB(A) when students were present. These levels are very similar to those measured in two recent surveys of classroom noise levels in Istanbul.[28,29] In another survey of noise during lectures in occupied university classrooms, Hodgson *et al.*[30] identify the contributions from particular sources to the overall noise levels, for example, ventilation noise (mean 40.9 dB), student-activity noise (mean 41.9 dB), and background level (mean 44.4 dB).

In the UK there have been a few small surveys of classroom noise. A survey in 1989[33] found, in primary schools, a mean level of 47 dB(A) for empty classrooms, with a range of 35.0 to 64.2 dB(A), and, for occupied classrooms, a mean of 65 dB(A) and range of 47.5 to 81.3 dB(A). Another survey of seven UK primary school classrooms[34] measured background noise levels in empty classrooms from 35 to 45 dB L_{Aeq} and in occupied classrooms with the children talking and working from 58 to 72 dB L_{Aeq}. These studies agree with the studies elsewhere in finding that 35 dB(A) is the lowest level likely to occur in an empty classroom, but that in an occupied classroom the levels are significantly higher.

More recently Mackenzie and Airey[25,35] measured average background noise levels of 44.7 dB(A) in empty classrooms, and 55.5 and 77.3 dB(A) when the children were silent and working, respectively. Other recent studies have found preschool children exposed to levels of 75 dB(A) in the classroom[24] and older children working in levels of 58 to 69 dB(A) during mathematics classes.[36]

In order to reduce the effects of noise on school children it is essential to determine the types of sound in a classroom in order to apply the most appropriate noise control methods. Thus, it is necessary to understand what determines classroom noise levels, that is, the noise sources in the classroom and, in particular, the impact of external noise upon the internal environment. The previous surveys of classroom noise have shown a wide range in noise levels. However, the reported levels have in many cases been presented as single figure ratings in dB(A), with no explanation of whether these represent instantaneous or time-averaged sound levels, or whether they are maximum (e.g., L_{Amax}), ambient (e.g., L_{Aeq}) or background (e.g., L_{A90}) levels. Other measurements are presented in terms of L_{Aeq}, but without the time period being specified. Furthermore, there has been little reported examination of the factors that determine classroom levels, for example occupancy of the classroom, activity of the children, or the presence of external noise. There has also been no discussion of the variation of noise in a classroom throughout the school day, or a comparison of classroom levels with other levels around a school. In particular, there has been no large-scale detailed study of levels in schools in an urban environment.

In the current study noise levels were measured outside 142 schools around London, and the types of noise sources present were examined, to give a general indication of the noise environment around schools in central London. Detailed measurements were also made, at different times to the external measurements, inside 16 of the schools to provide data on typical classroom noise levels of primary school children aged between 4 and 11, and to enable a comparison of internal levels with external levels. The variation of a number of noise parameters throughout the day in 140 classrooms was examined, and classroom noise levels were related to children's activities and age. Noise was also measured in other school locations and the effects of the age of the school buildings and of double glazing were also examined. External noise levels were compared with internal levels to determine the influence of external noise on the internal noise environment of the schools.

III. CHOICE OF SURVEY AREAS AND SCHOOLS

The overall aim of the study of which this survey was a part was to examine the influence of general environmental noise upon children in schools. It was therefore necessary to survey schools subject to a wide range of noise levels. Some

TABLE I. Means and standard deviations of external levels in each borough.

	$L_{Aeq,5\ min}$		$L_{A10,5\ min}$		$L_{A90,5\ min}$		$L_{A99,5\ min}$		$L_{Amax,5\ min}$		$L_{Amin,5\ min}$	
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	Mean	sd	Mean	sd
Haringey	57.4	8.8	59.4	9.0	49.2	7.7	47.0	7.4	70.5	10.5	46.0	7.5
Islington	56.2	9.4	58.4	9.9	46.5	9.3	44.3	9.2	68.3	17.0	41.3	12.4
Lambeth	58.9	7.4	61.2	7.7	50.2	8.2	47.8	8.2	72.0	9.0	47.0	8.3

schools in London are subject to high levels of noise being located close to main roads or industrial areas, whereas other London schools are sheltered from road traffic noise by surrounding buildings. It was therefore decided to focus on areas of London to provide a range of external noise levels and sources. As there already exists a considerable body of research studying the effects of aircraft noise on children,[10–15,37] it was decided that areas of London in which aircraft are the dominant environmental noise source (that is, areas to the west of London) should be excluded from the current survey. The choice of areas in which to carry out the survey was further determined by an examination of demographic data in boroughs across London, so as to choose areas that were typical of the demography of London as a whole. Data on educational attainments of primary schools across London were also examined to ensure that the schools in the areas of London selected for the surveys reflected the academic performance of primary schools across London.

Three London boroughs (Haringey, Islington, and Lambeth) were chosen according to the criteria described above, and so as to include schools within both inner London, that is, within approximately one mile of central London, and outer London, that is, approximately five miles from central London. Noise levels were measured outside every primary school in Haringey ($n=53$) and Islington ($n=50$), and outside a majority of schools in Lambeth ($n=39$). Detailed noise surveys were carried out in 16 of the schools in one outer and one inner borough (Haringey and Islington).

Primary schools in London generally fall into one of two types. Many are in large Victorian buildings, built in the latter half of the 19th century. These are brick buildings, often two or three stories high, with large windows and high ceilings, many of which are surrounded by large grounds and playgrounds, separating them from the nearest road. Other schools are in more modern buildings, typically built during the 1960s or 1970s. These tend to be low rise buildings with many windows. The locations of the schools in relation to noise sources vary; some are adjacent to busy roads; others are set back from the road, separated from the curbside by playgrounds; and many are set away from main roads on side streets.

Some of the schools have windows consisting of a single pane of glass throughout (''single glazing'') while others have double paned windows (''double glazing''), or secondary panes fitted internally to the original windows to increase sound insulation (''secondary glazing'').

IV. EXTERNAL NOISE SURVEY

A. Measurement method

Five minute samples of noise were measured outside each school using a Bruel and Kjaer handheld sound level meter, Type 2236. For security reasons measurements were made off the school premises, where possible outside the noisiest façade, at the curbside of the nearest road. In most cases this was at approximately 1 m from the nearside lane of traffic. For many schools the measurement position was at approximately 4 m from the school façade. For consistency, where measurements were at other distances from the traffic or from a school, the appropriate distance correction was applied to give the corresponding level 4 m from the façade.

The 5 min measurement period was chosen to be typical of the school day when the children would be working in the classroom. Thus, rush hour periods, times when children were arriving at or being collected from school, lunch hours, and times when children were outside in the school playground were avoided. Furthermore, it was felt that the noise climate during the measurement was typical of the noise environment of the area.

The environmental noise parameters $L_{Aeq,5\ min}$, $L_{A10,5\ min}$, $L_{A90,5\ min}$, $L_{A99,5\ min}$, $L_{Amax,5\ min}$, and $L_{Amin,5\ min}$ were recorded at each site. These parameters give an indication of the ambient ($L_{Aeq,5\ min}$), background ($L_{A90,5\ min}$), and underlying ($L_{A99,5\ min}$) noise characteristics of the local environment. $L_{A10,5\ min}$ was included in the measurements, as it gives an indication of the higher noise levels, and is used in the UK for the assessment of road traffic noise. $L_{Amax,5\ min}$ and $L_{Amin,5\ min}$ were also measured to show typical maximum and minimum levels to which schools may be exposed during the school day. All these parameters were subsequently compared with the internal levels measured, in an attempt to determine the characteristics of external environmental noise that affect internal levels. They were also compared with the results of a questionnaire survey of children's and teachers' attitudes to noise,[5] to assess those aspects of noise that affect annoyance.

In addition to noise levels, during the 5 min measurement period the noise sources heard by the researchers were noted.

B. Results of external noise survey

1. Measured noise levels

The means and standard deviations of the measured parameters for each borough are shown in Table I.

TABLE II. Means, standard deviations, and ranges of external levels [dB(A)].

	$L_{Aeq,5\ min}$	$L_{A10,5\ min}$	$L_{A90,5\ min}$	$L_{A99,5\ min}$	$L_{Amax,5\ min}$	$L_{Amin,5\ min}$
Mean	57.4	59.6	48.5	46.3	70.1	44.6
sd	8.7	9.0	8.6	8.6	12.9	10.0
Range	31–78	32–81	25–71	21–68	42–93	20–67

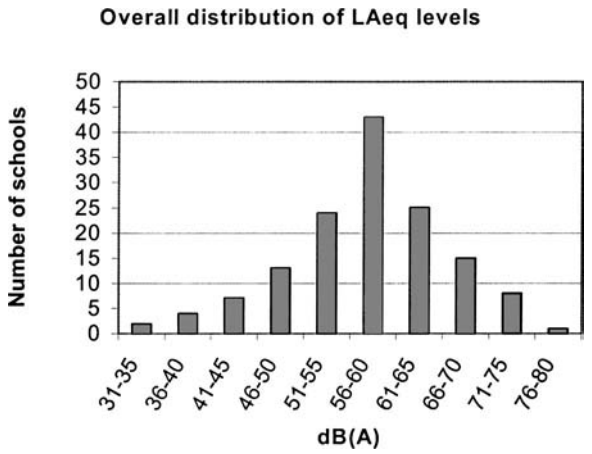


FIG. 1. Distribution of external $L_{Aeq,5\ min}$ levels.

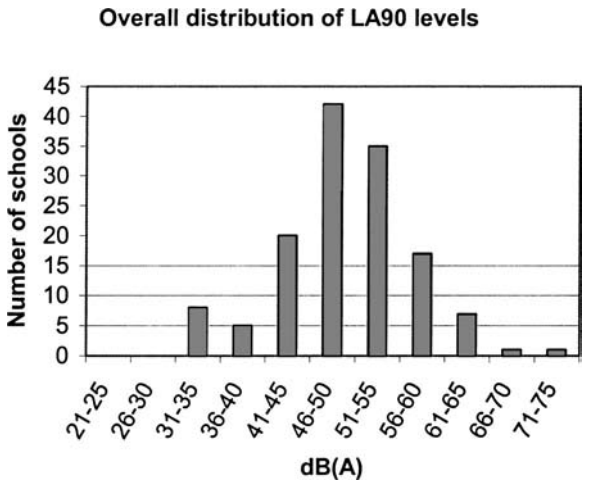


FIG. 2. Distribution of external $L_{A90,5\ min}$ levels.

It can be seen that the values of all parameters are similar across the three boroughs, although the means of all parameters in Islington are between 1 and 5 dB(A) below those in the other two boroughs. Many schools in Islington are located on side streets, which are particularly quiet, being sheltered from main roads. Also, many of these inner city schools are in large Victorian buildings that typically are surrounded by large grounds and playgrounds, separating them from the nearest road. The subjective impression formed when carrying out the external noise survey was that noise levels outside several schools in this area were unrepresentative of (normally lower than) the general environmental noise climate in the area. It can also be seen that the standard deviations of all parameters for Islington are greater than for the other two boroughs; this again reflects the fact that many Islington schools are in very quiet surroundings on side streets, whereas other schools in the borough are on main roads and therefore exposed to high levels of road traffic noise. Thus the distribution of noise levels outside schools in this borough may not be typical of London as a whole.

Considering all schools together, the means, standard deviations, and ranges of the measured parameters are shown in Table II. There is a relatively small [11 dB(A)] difference between L_{A10} and L_{A90} levels; this is to be expected of levels measured during the day in an urban environment. It can be seen that for most parameters the standard deviation is approximately 9 dB(A). The greatest variation in levels occurs for the L_{Amax} levels, with a high standard deviation of approximately 13 dB(A). The L_{Amax} measured during a 5 min period will reflect the occurrence of individual events with noise levels higher than the ambient noise. This parameter would therefore be expected to demonstrate the widest variation of all parameters.

The wide range of levels occurring is illustrated in Figs. 1–3, which show the distributions of the parameters $L_{Aeq,5\ min}$, $L_{A90,5\ min}$, and $L_{Amax,5\ min}$ measured outside all schools. The figures show that the most commonly occurring ambient levels are in the range of 56–60 dB L_{Aeq}, and background levels from 46 to 50 dB L_{A90}.

2. Sources of noise

During the external noise survey observations were made of the noise sources heard outside each school during the 5 min sampling period. Figure 4 shows the incidences of the most commonly occurring sources. It can be seen that, as would be expected in an urban environment, the most common source of noise was road traffic, principally cars, which was heard outside 86% of schools. Sirens were heard at surprisingly few schools, although they are commonly regarded as a regular feature of the London noise environment, and reported as being frequently heard by teachers and children.[5]

It can also be seen that, although west London boroughs subject to predominantly aircraft noise were excluded from the study, aircraft were still heard at over 50% of the schools.

V. INTERNAL NOISE SURVEY

Detailed internal noise surveys were carried out in 16 of the schools measured in the external survey, eight in the outer London borough (Haringey) and eight in the inner borough of Islington. The schools were chosen to give examples from across the range of external noise levels measured, the external L_{Aeq} levels outside the 16 schools ranging from 49 to 75 dB(A). As with the external surveys it was found that levels were consistent across the two boroughs; for this reason the boroughs have not been considered separately, all 16 schools are considered together.

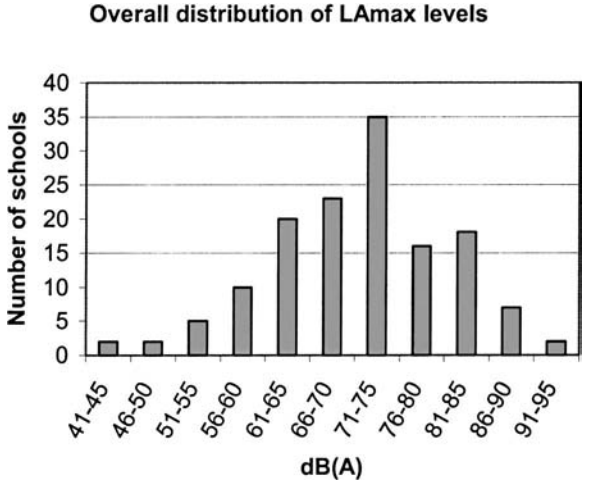


FIG. 3. Distribution of external $L_{Amax,5\ min}$ levels.

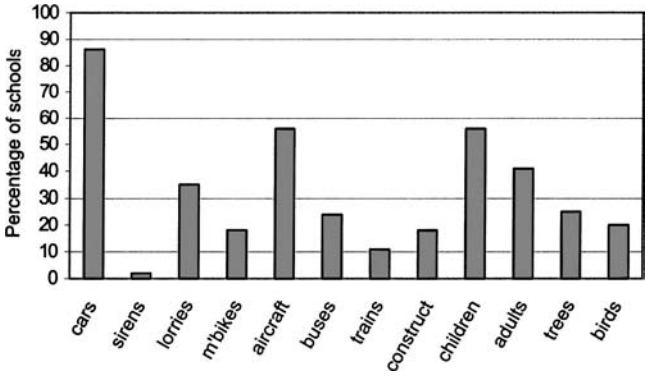


FIG. 4. Incidence of commonly occurring noise sources outside schools.

Measurements were made in approximately 200 school locations, including 110 occupied classrooms, 30 empty classrooms and 50 other school locations. Approximately half of the school buildings were Victorian and half dated from the latter half of the 20th century. Around 50% of the schools were single glazed, the remainder having some form of double or secondary glazing. At the time of the surveys, all windows were shut. Ventilation/heating systems were in operation in some schools at the time of measurement.

In addition to noise surveys in each school, a detailed classroom observation was carried out to record the activities the children were undertaking in the classrooms at the times of the measurements and to note any noise sources that were particularly obvious. It was therefore possible to relate the measured noise levels to classroom activities, as well as to the number of children in the class and the age of the children. The effects of the age of the school buildings and of double glazing were also considered.

A. Method

When measuring noise in schools various important practical issues need to be taken into account. To obtain representative data in occupied classrooms care must be taken not to disrupt the work of the teachers, and not to disturb or distract the children. Safety considerations are also important when measuring noise in the presence of young children. Therefore, prior to the main survey, pilot studies using various measurement techniques, and classroom observation, were carried out to determine the most appropriate form of noise measurement in classrooms.[38]

As a result of the pilot study it was decided that the most appropriate technique for the measurement of noise in primary school was the use of a handheld sound level meter. Short (2 min) samples of noise measured in this way in classrooms and other school locations give a good indication of the fluctuation in noise during the day in a classroom and of the variation in level throughout a school. Furthermore, using this method in an occupied classroom does not appear to interfere with the teaching or affect the children's concentration.

During the pilot study, continual noise monitoring during a morning or afternoon classroom session showed that the fluctuations of all noise parameters with time were very similar, and that the relative values of all parameters were approximately constant. For this reason the following discussion is confined to L_{Aeq} and L_{A90} levels only.

In each school, 2 min measurements of L_{Aeq} and L_{A90} were made during lessons in classrooms and other occupied and unoccupied spaces around the school, such as assembly halls, foyers, stairs and corridors, and empty classrooms. For each measurement the time, the current activity and occupancy (number of teachers and children) of the space, and the occurrence of any noticeable noise events, internal or external, were noted.

B. Results of classroom observation

1. Occupied classrooms

During the noise surveys in occupied classrooms the ages of the children, details of the classroom activity and occupancy (numbers of teachers, other adults, and children) were noted, together with the types of noise that were present. The subjective impression formed was that during lessons it was normally not possible to hear external noise or internal background noise from building services or classroom equipment. The classroom noise appeared to be dominated by the noise of the children themselves, and dependent upon the particular classroom activity that was being carried out. The lack of intrusion of external noise during the majority of activities was confirmed by subsequent correlation analysis of internal and external levels (see Sec. V).

Subjectively it was found that, in general, classroom sessions could be broken down into six distinct activities, each with a characteristic noise level resulting from all the sources related to the activity, including the teacher's voice.

The six activities identified were as follows:

(1) Activity 1: Children sitting at tables doing silent reading or tests.
(2) Activity 2: Children sitting at tables or on the floor, with one person (teacher or child) speaking at any one time.
(3) Activity 3: Children sitting at tables working individually, with some talking.
(4) Activity 4: Children working individually, moving around the classroom, with some talking.
(5) Activity 5: Children working in groups, sitting at tables, with some talking.

TABLE III. Average L_{Aeq} and L_{A90} levels for different age groups.

	Class/Age							
	Nursery (3–4)	Reception (4–5)	Year 1 (5–6)	Year 2 (6–7)	Year 3 (7–8)	Year 4 (8–9)	Year 5 (9–10)	Year 6 (10–11)
L_{Aeq}	71.9	73.9	74.3	66.3	68.9	69.6	73.2	71.2
L_{A90}	57.3	62.3	61.0	51.3	52.5	49.8	53.8	52.9

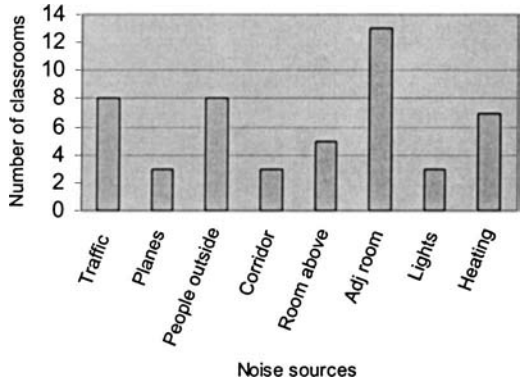


FIG. 5. Incidence of commonly occurring noise sources in empty classrooms.

(6) Activity 6: Children working in groups, moving around the classroom, with some talking.

2. Unoccupied classrooms

In all except one of the empty classrooms surveyed the doors were closed during the measurement period. The audible noise included external noise from road traffic, planes and people (adults and children), and internal noise from heating or ventilation and lighting systems. Noise was also heard from other parts of the school including the corridors outside the classrooms, adjacent or nearby teaching spaces and classrooms above the rooms being measured. The occurrences of the different sources heard in the empty classrooms are shown in Fig. 5.

C. Results of internal noise survey

In total, in the 16 schools over 220 measurements were made in 110 occupied classrooms, 30 empty classrooms, and 50 other locations, including occupied and empty assembly halls and corridors. The L_{Aeq} and L_{A90} levels measured have been analyzed as follows. For each school the data were averaged according to year group, classroom activity, and type of space. The data for individual schools were then combined to give overall average figures. The relationships between each of the following factors and noise levels have been examined: numbers of children in the classroom; ages of children; type of glazing (single or double); and age of the school buildings.

1. Variation of noise with number of children

The number of children in classes measured during the surveys varied from 7 to 32, although there were only one or two samples for numbers lower than 18. Figure 6 shows the relationships between average L_{Aeq} and L_{A90} levels corresponding to class sizes of greater than 18. These levels are the averages of all samples for the relevant number of children. There is significant correlation between class size and ambient L_{Aeq} level ($r=0.669$, $p<0.01$) and background level L_{A90} ($r=0.566$, $p<0.05$).

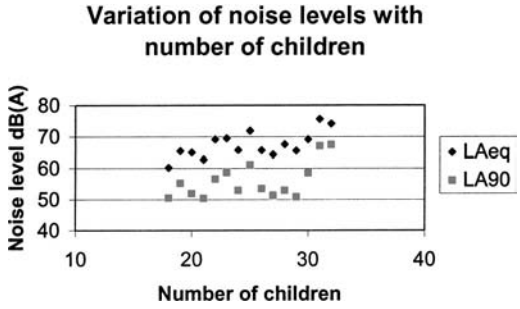


FIG. 6. Relationship between classroom L_{Aeq} and L_{A90} levels and number of children.

2. Variation of noise with age

The L_{Aeq} and L_{A90} levels corresponding to different year groups, averaged over all activities, are shown in Table III. It was found that, in several of the schools, there was a general trend for the noise levels to decrease as the age of the children increases. There is anecdotal evidence among teachers that this is the case and that nursery and reception classes can be particularly noisy. Picard and Bradley,[38] in reviewing published data on noise levels in schools, found a general reduction in classroom noise levels with increasing age. However, the occurrence of high noise levels in some Year 5 and Year 6 classes meant that the decreasing pattern was not repeated in all schools in the current survey, and is not reflected in the average levels for the 16 schools. Furthermore, there is no evidence of noise reduction with age if the levels for each activity are broken down into the different age groups.

3. Variation of noise with activity

The average L_{Aeq} and L_{A90} levels for each of the six classroom activities listed above are shown in Table IV.

It can be seen that the noise levels increase with activity number, as would be expected, given the nature of the work and communications involved in the activities defined above. There is a difference of 20 dB(A), on average, between the quietest and noisiest classroom activities. When the children are engaged in the quietest activity, such as working in silence, doing a test or silent reading, the average ambient noise level is 56 dB L_{Aeq}. It is interesting to note that this level agrees with that measured by Hodgson in a university

TABLE IV. Average L_{Aeq} and L_{A90} levels for different classroom activities.

	Activity					
	Activity 1 Silent reading/test	Activity 2 1 person speaking	Activity 3 Individual work	Activity 4 Individual work and movement	Activity 5 Group work	Activity 6 Group work and movement
L_{Aeq}	56.3	61.2	64.7	72.2	72.9	76.8
L_{A90}	42.4	45.8	52.1	59.6	58.6	63.9

TABLE V. Average L_{Aeq} and L_{A90} levels in various school locations.

	Space				
	Occupied teaching space	Unoccupied classrooms	Corridor/ foyer/ stairs	Occupied hall	Unoccupied hall
L_{Aeq}	72.1	47.0	58.1	73.4	53.2
L_{A90}	54.1	36.9	44.6	55.1	44.3

classroom with students present, but silent.[32] The noisiest activity, Activity 6, has an average level of 77 dB L_{Aeq}. The measured levels also agree closely with those of MacKenzie,[25] who recorded an average background level of 55.5 dB(A) in acoustically untreated classrooms when pupils were silent, and an average level of 77.3 dB(A) when pupils were working.

4. Variation of noise within a school

Table V shows the averaged L_{Aeq} and L_{A90} levels measured within each school in the following types of space: occupied classrooms, unoccupied classrooms, corridors and foyer areas, occupied assembly halls, and empty assembly halls.

The average L_{Aeq} of occupied classrooms is 72 dB(A). It can therefore be assumed that this represents a child's noise exposure during a school day. The average L_{A90} of occupied classrooms is 54 dB(A). The areas with the lowest noise levels, in terms of both L_{Aeq} and L_{A90} levels, are empty classrooms with an average L_{Aeq} of 47 dB(A) and L_{A90} of 37 dB(A). There is thus a difference of 25 dB L_{Aeq}, on average, between the ''noisiest'' and ''quietest'' areas in a school, that is, between occupied and unoccupied classrooms. It can also be seen that the difference in noise level between an empty classroom and a classroom with children being ''silent'' is 9 dB L_{Aeq}. This is similar to the difference found by MacKenzie[25] in acoustically untreated primary school classrooms, where the average unoccupied level was 44.7 dB(A) and the average level with pupils silent was 55.5 dB(A). Hodgson[32] refers to the noise when pupils are silent as ''student-generated background noise,'' which includes noise such as the movement of chairs and the rustling of paper, but not voices.

If the L_{Aeq} of empty classrooms is regarded as the ''background'' level, as defined in the World Health Organization and other guidelines,[2–4] then the averaged level measured here is 12 dB(A) higher than the guideline value.

5. Noise levels in empty classrooms

It was not possible from the results of the noise survey to identify the contribution of particular sources to the overall noise level, as reported by Hodgson *et al.*[30] Four of the seven rooms where heating/ventilation noise was heard were in the same school and an examination of all the sound levels measured shows that the average levels for unoccupied classrooms in this school are the highest average levels measured out of the 16 schools (51.8 dB L_{Aeq} and 49.0 dB L_{A90}, compared with the average levels for empty classrooms of 47.0 dB L_{Aeq} and 36.9 dB L_{A90}). However, the occupied classroom levels for this school are not higher than those for other schools (70.1 dB L_{Aeq} and 54.0 dB L_{A90}, compared with the average levels of 72.1.0 dB L_{Aeq} and 54.1 dB L_{A90}), suggesting that ventilation system noise may not increase the noise levels in occupied classrooms.

6. Effects of age of school buildings

Of the 16 schools surveyed, 6 were in Victorian buildings and 7 in 20th century buildings dating from the 1960s or later. The remaining three schools were housed in a mixture of Victorian and modern buildings. The space average noise levels of the Victorian schools and of the modern schools have been compared (the three schools in ''mixed'' buildings have been omitted from this analysis) to see if there are any differences between noise levels in ''old'' and ''new'' schools. The average levels of the two types of school are shown in Table VI. In unoccupied classrooms and assembly halls the background and underlying noise levels are the same for the two types of school. However, the levels in occupied spaces and in circulation spaces are slightly higher in the Victorian schools than in the more modern schools. This could be because room volumes in Victorian schools tend to be greater than those in modern buildings, with a corresponding increase in the amount of reflective surface area, so that the reverberant sound level may be higher in general in the Victorian schools. However, the sample size is too small for definite conclusions regarding the effect of the age of school buildings to be made.

7. Effects of glazing

Here 6 of the 16 schools in the internal survey had windows of single pane glass while the other 10 had some form of secondary or double glazing. Insufficient data were available for any quantification of the sound insulation of the schools' facades, or of the glazing. A measurement of the sound insulation was beyond the scope of this survey, and as the schools in the survey were all between 30 and 130 years old, no specifications of the building materials used were

TABLE VI. Average noise levels in Victorian and modern schools.

	Occupied teaching space		Unoccupied classrooms		Corridors/ foyers/stairs		Occupied assembly hall		Unoccupied assembly hall	
	Vict[n]	Mod	Vict[n]	Mod	Vict[n]	Mod	Vict[n]	Mod	Vict[n]	Mod
L_{Aeq}	72.0	71.5	46.0	49.2	61.3	55.0	75.7	70.2	54.4	51.3
L_{A90}	56.2	52.5	38.2	38.5	46.2	43.3	58.7	50.2	44.8	45.0

TABLE VII. Average noise levels in schools with single and secondary/double glazing.

	Occupied teaching space		Unoccupied classrooms		Corridors/foyers/stairs		Occupied assembly hall		Unoccupied assembly hall	
	Single glazing	Secondary glazing	Single glazing	Secondary glazing	Single glazing	Secondary glazing	Single glazing	Secondary glazing	Single glazing	Secondary glazing
L_{Aeq}	71.4	71.8	47.3	47.4	60.6	55.6	70.9	74.4	54.7	49.7
L_{A90}	54.6	53.8	38.1	36.0	46.4	43.7	58.2	53.3	46.3	41.7

available. However, an attempt was made to examine the effectiveness of secondary/double glazing in schools by comparing the average space noise levels of the group of single glazed schools with those of the group of ten schools with secondary/double glazing. The average levels of the two groups are shown in Table VII. There is no repeated pattern to the differences between the groups: the ambient, background and underlying levels are very similar for occupied and unoccupied classrooms for both groups. However, it can be seen from Table VII that the background and underlying levels are slightly less, by between 1 and 5 dB(A), for the group of schools with secondary glazing than for the single glazed schools. Again the sample size is too small for definite conclusions to be drawn.

VI. RELATIONSHIP BETWEEN INTERNAL AND EXTERNAL NOISE

As stated earlier, the subjective impression formed during the survey of internal classroom noise was that the noise was dominated by the sound of the children's activities and was therefore not dependent upon the external noise environment.

To examine this further, for the 16 schools in the internal noise survey the internal L_{Aeq} and L_{A90} levels, classified by activity, age, and space, were correlated with all the measured external noise parameters. It is recognized that this provides a relatively crude examination of the effects of external noise on the indoor noise environment, given that it was only possible to compare averaged internal and external noise levels, indoor and outdoor levels not having been measured simultaneously. However, statistically significant correlations were found between the L_{Aeq} for Activity 1 (''quiet,'' test conditions) and the external L_{Amin}, L_{A99}, and L_{A90} levels. The correlation coefficients were high: 0.962, 0.975, and 0.960, respectively, all statistically significant at the 1% level. Thus, it would appear that when children are engaged in quiet activities in the classroom the ambient classroom level is closely related to the background and underlying levels outside. This is consistent with the results of the questionnaire survey of over 2000 children and their teachers that was carried out in the borough of Haringey in parallel with the noise surveys. Children, particularly in the older age group (11 years), reported being able to hear a variety of external noise sources while in the classroom, and over 90% of the teachers questioned felt that noise affected the pupils' concentration.[5]

An attempt was made to further examine the effectiveness of secondary/double glazing by correlating internal and external levels for each of the groups of single glazed schools and those with secondary/double glazing. There were not sufficient data to obtain meaningful results for the set of single glazed schools. For the secondary/double glazed group, although there were only four schools for which there was relevant data, there were still significant positive correlations between Activity 1 L_{Aeq} levels and external L_{Amin}, L_{A99}, and L_{A90} levels (L_{Amin}: $r=0.914$, $p<0.05$; L_{A99}: $r=0.949$, $p<0.05$; L_{A90}: $r=0.995$, $p<0.01$). This suggests that the secondary/double glazing is ineffective in these cases. This may be due to the specification or fitting of the glazing, or simply to the fact that no particularly ''noisy'' events occurred outside these schools at the times of the measurements. Alternatively it could be because the predominant external noise is road traffic noise which tends to be mainly low frequency in character, and therefore more difficult to attenuate by glazing. However, there is insufficient data to draw any firm conclusions on the effectiveness of double or secondary glazing in schools.

VII. CONCLUSIONS

A survey of noise levels outside 142 primary schools in 3 London boroughs has shown that the average L_{Aeq}, measured over a typical 5 min period during the school day, is approximately 57 dB(A). However, there was a wide range of levels for all the parameters measured, with some schools in supposedly ''noisy'' areas being exposed to relatively low levels of noise. This suggests that in work concerning environmental noise exposure at school it is necessary to measure the noise rather than rely on noise contours or noise maps to give an indication of a school's noise exposure level.

The predominant noise source outside the London schools surveyed was road traffic, in particular, cars, which could be heard outside 86% of the schools. Although schools where aircraft noise dominates the noise environment were deliberately excluded from this study, aircraft could be heard at over 50% of the schools surveyed.

The noise inside classrooms is, in general, dominated by the noise of children and depends upon the particular classroom activity in which they are engaged, there being a range of approximately 20 dB(A) between the quietest and noisiest activity. The age of the children was not closely related to noise levels, but there were significant relationships between the number of children in a classroom and the ambient and background classroom noise levels. The average L_{Aeq} of occupied teaching spaces, which could be assumed to be the average exposure for a child at school, was found to be 72 dB L_{Aeq}.

Subjectively external noise appeared to have little effect on the internal noise environment. Correlation analysis con-

firmed that this is the case for the majority of classroom activities. However, when children are engaged in a quiet activity such as silent reading or doing a test, then the noise level in the classroom is closely related to the background and underlying levels outside. Thus children may be distracted by the noise and their concentration affected at times when they are working in silent conditions.

The survey has shown that the presence of pupils, even when silent, increases the noise level in a classroom. The appropriate measurement to compare with guideline values is therefore the average L_{Aeq} in *unoccupied* classrooms. In this survey this level was 47 dB(A), which is 12 dB(A) above the level of 35 dB(A) recommended by published guidelines.[2–4]

The levels measured relate to schools in an urban area. It is reasonable to assume that schools in rural areas, especially those away from main roads and airports, would have lower noise exposures. However, as the internal classroom noise depends on classroom activity, it could be assumed that internal levels in other schools would be similar to those in schools in urban areas. Further investigation is needed to examine noise levels in schools in suburban and rural areas for comparison with urban schools.

ACKNOWLEDGMENTS

The authors would like to thank the UK Department of Health and the Department of the Environment, Food and Regional Affairs for financing this project, and the schools, children, and teachers who took part in the research. Thanks are also due to Rebecca Asker for collection of the noise and classroom observation data.

[1] M. Vallet and Z. Karabiber, ''Some European policies regarding acoustical comfort in educational buildings,'' Noise Control Eng. J. **50**, 58–62 (2002).
[2] World Health Organization, *Guidelines for Community Noise*, ⟨http://www.who.int/peh/⟩, 1999.
[3] ANSI/ASA S12.60-2002, *Acoustical Performance Criteria, Design Requirements and Guidelines for Schools*, American National Standards Institute, 2002.
[4] Department for Education and Skills, *Building Bulletin 93: Acoustic Design for Schools*, The Stationary Office, London, 2003 (http://www.teachernet.gov.uk/acoustics/).
[5] J. E. Dockrell and B. M. Shield, ''Children's perceptions of their acoustic environment at home and at school,'' submitted to J. Am. Acoust Soc.
[6] B. Berglund and T. Lindvall, *Community Noise*, Document prepared for the World Health Organization, Archives of the Center for Sensory Research 2(1), Stockholm, University and Karolinska Institute, Sweden, 1995.
[7] ''The non-auditory effects of noise,'' Institute for Environment and Health, University of Leicester, UK, 1997.
[8] R. Hetu, C. Truchon-Gagnon, and S. A. Bilodeau, ''Problems of noise in school settings: a review of literature and the results of an exploratory study,'' J. Speech, Lang. Pathol. Audiol. **14**, 31–38 (1990).
[9] G. Evans and S. Lepore, ''Non-auditory effects of noise on children: a critical review,'' Children's Environments Quarterly **10**, 31–51 (1993).
[10] M. A. Crook and F. J. Langdon, ''The effects of aircraft noise in schools around London Airport,'' J. Sound Vib. **3**, 221–232 (1974).
[11] S. Cohen, G. W. Evans, D. S. Krantz, and D. Stokols, ''Physiological, motivational, and cognitive effects of aircraft noise on children. Moving from the laboratory to the field,'' Am. Psychol. **35**, 231–243 (1980).
[12] S. Cohen, G. W. Evans, D. S. Krantz, D. Stokols, and S. Kelly, ''Aircraft noise and children, longitudinal and cross sectional evidence on adaptation to noise and the effectiveness of noise abatement,'' J. Pers. Soc. Psychol. **40**, 331–345 (1981).
[13] S. Hygge, G. W. Evans, and M. Bullinger, ''A prospective study of some effects of aircraft noise on cognitive performance in schoolchildren,'' Psychol. Sci. **13**, 469–474 (2002).
[14] M. M. Haines, S. A. Stansfeld, J. Head, and R. F. S. Job, ''Multi-level modelling of aircraft noise on performance tests in schools around Heathrow Airport London,'' J. Epidemiol. Community Health **56**, 139–144 (2002).
[15] M. M. Haines, S. A. Stansfeld, R. F. S. Job, B. Berglund, and J. Head, ''Chronic aircraft noise exposure, stress responses, mental health and cognitive performance in school children,'' Psychol. Med. **31**, 265–277 (2001).
[16] A. L. Bronzaft and D. P. McCarthy, ''The effect of elevated train noise on reading ability,'' Environment Behaviour **7**, 517–527 (1975).
[17] A. L. Bronzaft, ''The effect of a noise abatement program on reading ability,'' J. Environmental Psychol. **1**, 215–222 (1981).
[18] J. W. Sargent, M. I. Gidmanm, M. A. Humphreys, and W. A. Utley, ''The disturbance caused to school teachers by noise,'' J. Sound Vib. **70**, 557–572 (1980).
[19] J. S. Lukas, R. B. DuPree, and J. W. Swing, ''Effect of noise on academic achievement and classroom behavior,'' FHWA/CA/DOHS-81/01. Federal Highways Works Agency, California Department of Health Services, Sacramento, 1981.
[20] S. Sanz, A. M. Garcia, and A. Garcia, ''Road traffic noise around schools: a risk for pupils' performance?,'' Int. Arch. Occup. Environ. Health **65**, 205–207 (1993).
[21] J. Romero and D. Lliso, ''Perception and acoustic conditions in secondary Spanish schools,'' *Proceedings of the 15th International Congress on Acoustics, Trondheim, Norway*, 1995, pp. 271–274.
[22] S. Hygge, ''Classroom experiments on the effects of aircraft, traffic, train, and verbal noise on long-term recall and recognition in children aged 12–14 years,'' in *Noise as a Public Health Problem*, Proceedings of the 6th International Congress, edited by M. Vallet, 1993, Vol. 2, pp. 531–534.
[23] F. S. Berg, J. C. Blair, and V. Benson, ''Classroom acoustics: The problem, impact, and solution,'' Speech, Language, Hearing Services in the Schools **27**, 16–20 (1996).
[24] L. Maxwell and G. Evans, ''The effects of noise on pre-school children's pre-reading skills,'' J. Environmental Psychol. **20**, 91–97 (2000).
[25] D. Mackenzie, ''Noise sources and levels in UK schools,'' *Proceedings of the International Symposium on Noise Control and Acoustics for Educational Buildings*, Istanbul, May 2000, Proceedings of the Turkish Acoustical Society, 2000, pp. 97–106.
[26] K. B. Green, B. S. Pasternack, and R. E. Shore, ''Effects of aircraft noise on reading ability of school-age children,'' Arch. Environ. Health **37**, 24–31 (1982).
[27] Y. Avsar and M. T. Gonullu, ''A map preparation for outdoor noises of educational buildings in Fatih district of Istanbul,'' in Ref. 25, pp. 69–76.
[28] S. Kurra, ''Results of a pilot study about teacher's annoyance relative to noise exposure in 3 high schools in Istanbul,'' in Ref. 28, pp. 47–56.
[29] E. Celik and Z. Karabiber, ''A pilot study on the ratio of schools and students affected from noise,'' in Ref. 25, pp. 119–128.
[30] M. Hodgson, R. Rempel, and S. Kennedy, ''Measurement and prediction of typical speech and background noise levels in university classrooms during lectures,'' J. Acoust. Soc. Am. **105**, 226–233 (1999).
[31] M. Picard and J. Bradley, ''Revisiting speech interference in classrooms,'' Audiology **40**, 221–244 (2001).
[32] M. Hodgson, ''UBC-Classroom acoustical survey,'' Can. Acoust. **22**, 3–10 (1994).
[33] A. Moodley, ''Acoustic conditions in mainstream classrooms,'' J. Br. Assoc. Teachers of the Deaf **13**, 48–54 (1989).
[34] B. Hay, ''A pilot study of classroom noise levels and teachers' reactions,'' Voice **4**, 127–134 (1995).
[35] S. Airey, ''A survey of acoustical standards in UK classrooms and their effect on pupils and teachers,'' Proc. Inst. Acoust. **20**, 14–21 (1998).
[36] P. Lundquist, K. Holmberg, and U. Landstrom, ''Annoyance and effects on work from environmental noise at school,'' Noise Health **2**, 39–46 (2000).
[37] K. Karami and S. Frost, ''Effects of aircraft noise on education in schools adjacent to Tehran Airport, Iran,'' Int. J. Environ Education Information **18**, 137–142 (1999).
[38] B. Shield, R. Jeffery, J. Dockrell, and I. Tachmatzidis, ''A noise survey of primary schools in London,'' in Ref. 25, pp. 109–118.

A beam tracing method for interactive architectural acoustics

Thomas Funkhouser[a)]
Princeton University, Princeton, New Jersey 08540

Nicolas Tsingos
INRIA, Sophia-Antipolis, France

Ingrid Carlbom
Lucent Bell Laboratories, Murray Hill, New Jersey

Gary Elko and Mohan Sondhi
Avaya Labs, Basking Ridge, New Jersey 07920

James E. West
The John Hopkins University, Baltimore, Maryland 21218

Gopal Pingali
IBM TJ Watson Research Center, Hawthorne, New York 10532

Patrick Min and Addy Ngan
Princeton University, Princeton, New Jersey 08540

(Received 9 May 2002; revised 11 April 2003; accepted 25 August 2003)

A difficult challenge in geometrical acoustic modeling is computing propagation paths from sound sources to receivers fast enough for interactive applications. This paper describes a beam tracing method that enables interactive updates of propagation paths from a stationary source to a moving receiver in large building interiors. During a precomputation phase, convex polyhedral beams traced from the location of each sound source are stored in a ''beam tree'' representing the regions of space reachable by potential sequences of transmissions, diffractions, and specular reflections at surfaces of a 3D polygonal model. Then, during an interactive phase, the precomputed beam tree(s) are used to generate propagation paths from the source(s) to any receiver location at interactive rates. The key features of this beam tracing method are (1) it scales to support large building environments, (2) it models propagation due to edge diffraction, (3) it finds all propagation paths up to a given termination criterion without exhaustive search or risk of under-sampling, and (4) it updates propagation paths at interactive rates. The method has been demonstrated to work effectively in interactive acoustic design and virtual walkthrough applications. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1641020]

PACS numbers: 43.55.Ka, 43.58.Ta [VWS] Pages: 739–756

I. INTRODUCTION

Geometric acoustic modeling tools are commonly used for design and simulation of 3D architectural environments. For example, architects use CAD tools to evaluate the acoustic properties of proposed auditorium designs, factory planners predict the sound levels at different positions on factory floors, and audio engineers optimize arrangements of loudspeakers. Acoustic modeling can also be useful for providing spatialized sound effects in interactive virtual environment systems.[1,2]

One major challenge in geometric acoustic modeling is accurate and efficient computation of propagation paths.[3] As sound travels from source to receiver via a multitude of paths containing reflections, transmissions, and diffractions (see Fig. 1), accurate simulation is extremely compute intensive. Most prior systems for geometric acoustic modeling have been based on image source methods[4,5] and/or ray tracing,[6] and therefore they do not generally scale well to support large 3D environments, and/or they fail to find all significant propagation paths containing edge diffractions. These systems generally execute in ''batch'' mode, taking several seconds or minutes to update the acoustic model for a change of the source location, receiver location, or acoustical properites of the environment,[7] and they allow visual inspection of propagation paths only for a small set of prespecified source and receiver locations.

In this paper, we describe a beam tracing method that computes early propagation paths incorporating specular reflection, transmission, and edge diffraction in large building interiors fast enough to be used for interactive applications. While different aspects of this method have appeared at computer graphics conferences,[8–10] this paper provides the first complete description of the proposed acoustic modeling system.

Briefly, our system executes as follows. During an offline precomputation, we construct a spatial subdivision in which 3D space is partitioned into convex polyhedra (cells). Later, for each sound source, we trace beams through the spatial subdivision constructing a ''beam tree'' data structure

a)Electronic mail: funk@cs.princeton.edu

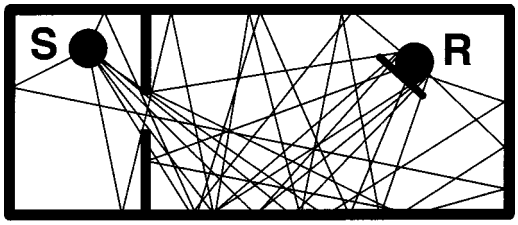


FIG. 1. Propagation paths.

encoding convex polyhedral regions of space reachable from the source by different sequences of scattering events. Then, during an interactive session, the beam trees are used to find propagation paths from the source and an arbitrary receiver location. The updates for each receiver are quick enough to be applied in an interactive acoustic design application. (For simplicity of exposition, our paper considers only propagation paths traced from sound sources to receivers. However, paths from receivers to sources could be computed just as easily—simply switch the terms ''source'' and ''receiver'' in the following text.)

The most important contribution of this paper is a method for precomputing data structures that encode potential sequences of surface scattering in a manner that enables interactive updates of propagation paths from a stationary source location to an arbitrarily moving receiver location. Our algorithms for construction and query of these data structures have the unique features that they scale well with increasing geometric complexity in densely occluded environments and that they generate propagation paths with any combination of transmission, specular reflection, and diffraction without risk of undersampling. We have incorporated these data structures and algorithms into a system that supports real-time auralization and visualization of large virtual environments.

The remainder of the paper is organized as follows. The next section reviews previous work in geometric acoustic modeling. Section III contains an overview of our system, with details of the spatial subdivision, beam tracing, path generation, and auralization methods appearing in Sec. IV. Section V contains experimental results. Applications, limitations, and topics for future work are discussed in Sec. VI. Finally, Sec. VII contains a brief conclusion.

II. PREVIOUS WORK

There have been decades of work in acoustic modeling of architectural environments, including several commercial systems for computer-aided design of concert halls (e.g., Refs. 11–13). Surveys can be found in Refs. 3 and 7.

Briefly, prior methods can be classified into two major types: (1) numerical solutions to the wave equation using finite/boundary element methods (FEM/BEM) and (2) high-frequency approximations based on geometrical propagation paths. In the latter case, image source methods, ray tracing, and beam tracing have been used to construct the sound propagation paths.

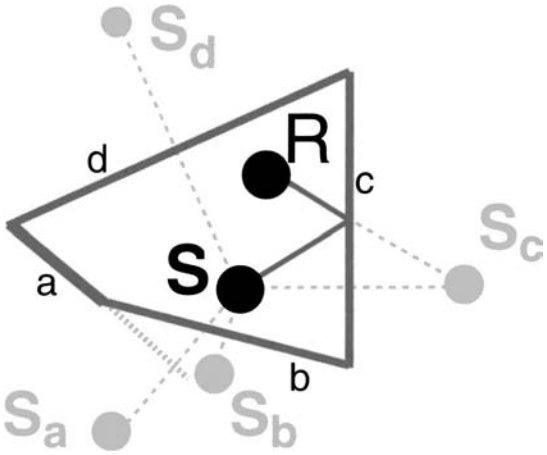


FIG. 2. Image source method.

A. Boundary element methods

Finite and boundary element methods solve the wave equation (and associated boundary conditions), subdividing space (and possibly time) into *elements*.[14–17] The wave equation is then expressed as a discrete set of linear equations for these elements. The boundary integral form of the wave equation (i.e., Green's or Helmoltz–Kirchhoff's equation) can be solved by subdividing only the boundaries of the environment and assuming the pressure (or particle velocity) is a linear combination of a finite number of basis functions on the elements. One can either impose that the wave equation is satisfied at a set of discrete points (collocation method) or ensure a global convergence criteria (Galerkin method). In the limit, finite element techniques provide an accurate solution to the wave equation. However, they are mainly used at low frequencies and for simple environments since the compute time and storage space increase dramatically with frequency.

Finite element techniques are also used to model *energy* transfer between surfaces. Such techniques have already been applied in acoustics,[18,19] as well as other fields,[20,21] and provide an efficient way of modeling diffuse global energy exchanges (i.e., where surfaces are lambertian reflectors). While they are well suited for computing energy decay characteristics in a given environment, energy exchange techniques do not allow direct reconstruction of an impulse response. Instead, they require the use of an underlying statistical model and a random phase assumption.[22] Moreover, most surfaces act primarily as specular or glossy reflectors for sound. Although extensions to nondiffuse environments have been proposed in computer graphics,[21,20] they are often time and memory consuming and not well suited to interactive applications.

B. Image source methods

Image source methods[4,5] compute specular reflection paths by considering *virtual sources* generated by mirroring the location of the audio source, S, over each polygonal surface of the environment (see Fig. 2). For each virtual source, S_i, a specular reflection path can be constructed by iterative intersection of a line segment from the source position to the

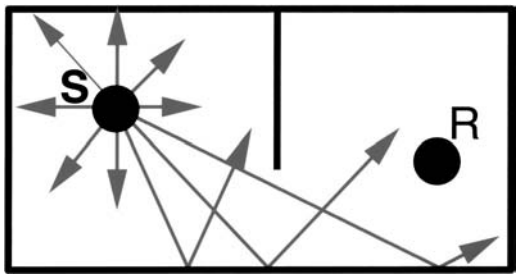


FIG. 3. Ray tracing method.

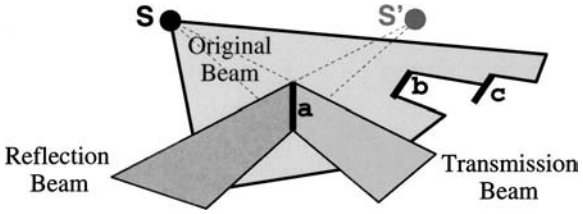


FIG. 4. Beam tracing method.

receiver position, *R*, with the reflecting surface planes (such a path is shown for virtual source S_c in Fig. 2). Specular reflection paths are computed up to any order by recursive generation of virtual sources.

The primary advantage of image source methods is their robustness. They guarantee that all specular paths up to a given order or reverberation time are found. However, the basic image source method models only specular reflection, and their expected computational complexity grows exponentially. In general, $O(n^r)$ virtual sources must be generated for r reflections in environments with n surface planes. Moreover, in all but the simplest environments (e.g., a box), complex validity/visibility checks must be performed for each of the $O(n^r)$ virtual sources since not all of the virtual sources represent physically realizable specular reflection paths.[5] For instance, a virtual source generated by reflection over the nonreflective side of a surface is "invalid."[5] Likewise, a virtual source whose reflection is blocked by another surface in the environment or intersects a point on a surface's plane which is outside the surface's boundary (e.g., S_a in Fig. 2) is "invisible."[5] During recursive generation of virtual sources, descendents of invalid virtual sources can be ignored. However, descendents of invisible virtual sources must still be considered, as higher-order reflections may generate visible virtual sources (consider mirroring S_a over surface *d*). Due to the computational demands of $O(n^r)$ visibility checks, image source methods are practical for modeling only a few specular reflections in simple environments.[23]

C. Ray tracing methods

Ray tracing methods[6] find propagation paths between a source and receiver by generating rays emanating from the source position and following them through the environment until a set of rays has been found that reach the receiver (see Fig. 3).

The primary advantage of these methods is their simplicity. They depend only on ray–surface intersection calculations, which are relatively easy to implement and have computational complexity that grows sublinearly with the number of surfaces in the model. Another advantage is generality. As each ray–surface intersection is found, paths of specular reflection, diffuse reflection, diffraction, and refraction can be sampled,[24,25] thereby modeling arbitrary types of propagation, even for models with curved surfaces. The primary disadvantages of ray tracing methods stem from their discrete sampling of rays, which may lead to undersampling errors in predicted room responses.[26] For instance, the receiver position and diffracting edges are often approximated by volumes of space (in order to enable intersections with infinitely thin rays), which can lead to false hits and paths counted multiple times.[26] Moreover, important propagation paths may be missed by all samples. In order to minimize the likelihood of large errors, ray tracing systems often generate a large number of samples, which requires a large amount of computation. Another disadvantage of ray tracing is that the results are dependent on a particular receiver position, and thus these methods are not directly applicable in interactive applications where either the source or receiver can move.

D. Beam tracing methods

Beam tracing methods[27,28] classify propagation paths from a source by recursively tracing pyramidal beams (i.e., sets of rays) through the environment (see Fig. 4). Briefly, for each beam, polygons in the environment are considered for intersection with the beam in front-to-back visibility order (i.e., such that no polygon is considered until all others that at least partially occlude it have already been considered). As intersecting polygons are detected, the original beam is clipped to remove the shadow region, a transmission beam is constructed matching the shadow region, and a reflection beam is constructed by mirroring the transmission beam over the polygon's plane. This method has been used in a variety of applications, including acoustic modeling,[27,8,29–31] illumination,[32–35,28,36] visibility determination,[37–39] and radio propagation prediction.[40,41]

The primary advantage of beam tracing is that it leverages geometric coherence, since each beam represents an infinite number of potential ray paths emanating from the source location. It does not suffer from the sampling artifacts of ray tracing,[26] nor the overlap problems of cone tracing,[42,43] since the entire space of directions leaving the source can be covered by beams exactly. The disadvantage is that the geometric operations required to trace beams through a 3D model (i.e., intersection and clipping) are relatively complex, as each beam may be reflected and/or obstructed by several surfaces.

Some systems avoid the geometric complexity of beam tracing by approximating each beam by its medial axis ray for intersection and mirror operations,[44] possibly splitting rays as they diverge with distance.[45,46] In this case, the beam representation is useful only for modeling the distribution of rays/energy with distance and for avoiding large tolerances in ray–receiver intersection calculations. If beams are not clipped or split when they intersect more than one surface, significant propagation paths can be missed, and the computed acoustical field can be grossly approximated.

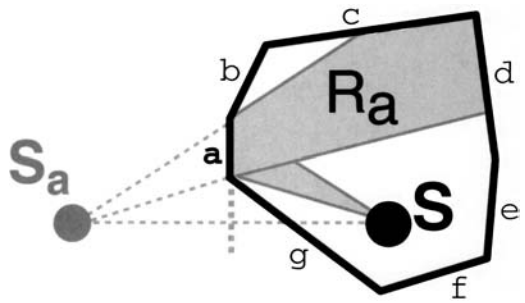


FIG. 5. Beam tracing culls invisible virtual sources.

III. OVERVIEW OF APPROACH

Our approach is based on polyhedral beam tracing. The strategy is to trace beams that decompose the space of rays into topologically distinct bundles corresponding to potential sequences of scattering events at surfaces of the 3D scene (*propagation sequences*), and then use them to guide efficient generation of propagation paths between a sound source and receiver in a later interactive phase. This approach has several advantages:

(i) **Efficient enumeration of propagation sequences**: Beam tracing provides a method for enumerating potential sequences of surface scattering events *without exhaustive search*, as in image source methods.[4,5] Since each beam describes the region of space containing all possible rays representing a particular sequence of scattering events, only surfaces intersecting the beam must be considered for further propagation. For instance, in Fig. 5, consider the virtual source S_a, which results from mirroring S over polygon a. The corresponding specular reflection beam, R_a, contains exactly the set of receiver points for which S_a is valid and visible. Similarly, R_a intersects exactly the set of polygons (c and d) for which second-order reflections are possible after specular reflection off polygon a. Other polygons (b, e, f, and g) need not be considered for higher-order propagation after specular reflection off a. As in this example, beam tracing can be used to prune the combinatorial search space of propagation sequences without resorting to sampling.

(ii) **Deterministic computation**: Beam tracing provides a method for finding potential sequences of diffracting edges and reflecting faces *without risk of errors due to under-sampling*, as in ray tracing. Since the entire 2D space of directions leaving the source can be partitioned so that every ray is in exactly one beam, beam tracing methods can guarantee finding every propagation path up to a specified termination criteria. Moreover, beams support well-defined intersections with points and edges, and thus beam tracing methods do not generate the systematic errors of ray tracing due to approximations made in intersecting infinitely thin rays with infinitely thin edges or infinitely small receiver points.[26]

(iii) **Geometric coherence**: Tracing beams can improve the efficiency of multiple ray intersection tests. In particular, once a beam has been traced along a certain sequence of surface intersections, generating a ray path from a source to a receiver following the same sequence requires only checking the ray path for intersections with the surfaces of the sequence, and the expensive computation of casting rays through a scene can be amortized over several ray paths. Beam tracing can be used not only to enumerate potential propagation sequences, but also to identify which elements of the scene can potentially be blockers for each sequence.[47,48] This information can be used to generate and check occlusion of sampled propagation paths quickly—i.e., in time proportional to the length of the sequence rather than the complexity of the scene.

(iv) **Progressive refinement**: Characteristics of the sound waves represented by beams can be used to guide priority-driven strategies.[9,49] For instance, estimates of the acoustic energy carried by different beams can be used to order beam tracing steps and to detect early termination criteria. This method is far more practical than precomputing a global visibility structure, such as the visibility skeleton,[50] which requires large amounts of compute time and storage, mostly for pairs of surfaces for which transport is insignificant. Instead, the proposed approach traces beams in priority order, finding propagation paths only as necessary for the required accuracy of the solution.

The main challenge of beam tracing is to develop methods that trace beams through 3D models robustly and efficiently and that generate propagation paths quickly. Although several data structures have been proposed to accelerate beam tracing computations, including ones based on binary space partitions,[27] cell adjacency graphs,[37,41,8,38,39] and layers of 2D triangulations,[40] no previous method models edge diffraction without sampling artifacts, and none provides interactive path updates in large 3D environments.

The key idea behind our method is to precompute and store spatial data structures that encode all possible sequences of surface and edges scattering of sound emanating from each audio source and then use these data structures to compute propagation paths to arbitrary observer viewpoints for real-time auralization during an interactive user session. Specifically, we use a precomputed polyhedral cell complex to accelerate beam tracing and a precomputed beam tree data structure to accelerate generation of propagation paths. The net result is that our method (1) scales to support large architectural environments, (2) models propagation due to edge diffraction, (3) finds all propagation paths up to a given termination criterion without exhaustive search or risk of undersampling, and (4) updates propagation paths at interactive rates. We use this system for interactive acoustic design of architectural environments.

IV. IMPLEMENTATION

Execution of our system proceeds in four distinct phases, as shown in Fig. 6. The first two phases execute

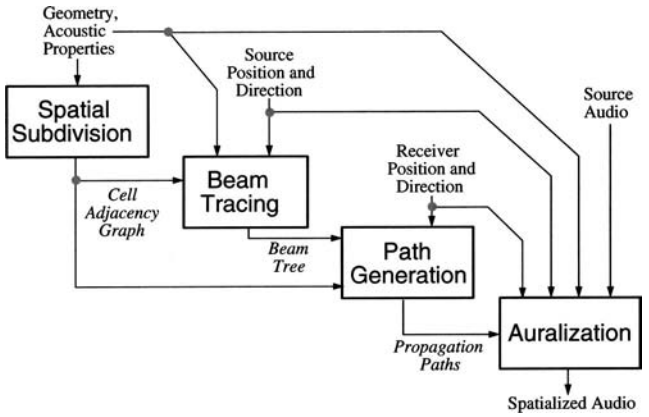


FIG. 6. System organization.

off-line, precomputing data structures for each stationary audio source, while the last two execute in real-time as a user moves the audio receiver interactively.

The result of each phase is shown in Fig. 7. First, during the *spatial subdivision phase*, we precompute spatial relationships inherent in the set of polygons describing the environment and represent them in a cell adjacency graph data structure that supports efficient traversals of space [Fig. 7(a)]. Second, during the *beam tracing phase*, we recursively follow beams of transmission, diffraction, and specular reflection through space for each audio source [Fig. 7(b)]. The output of the beam tracing phase is a beam tree data structure that explicitly encodes the region of space reachable by each sequence of reflection and transmission paths from each source point. Third, during the *path generation* phase, we compute propagation paths from each source to the receiver via lookup into the precomputed beam tree data structure as the receiver is moved under interactive user control [Fig. 7(c)]. Finally, during the *auralization phase*, we output a spatialized audio signal by convolving anechoic source signals with impulse response filters derived from the lengths, attenuations, and directions of the computed propagation paths [Fig. 7(d)]. The spatialized audio output is synchronized with real-time graphics output to provide an interactive audio/visual experience. The following subsections describe each of the four phases in detail.

A. Spatial subdivision

During the first preprocessing phase, we build a spatial subdivision representing a decomposition of 3D space and store it in a structure which we call a *winged-pair* representation. The goal of this phase is to partition space into convex polyhedral cells whose boundaries are aligned with polygons of the 3D input model and to encode cell adjacencies in a data structure enabling efficient traversals of 3D space during the later beam tracing phase.

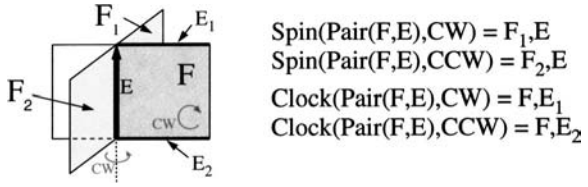


FIG. 8. Winged-pair structure.

The winged-pair data structure stores topological adjacencies in fixed-size records associated with vertices, edges, faces, cells, and face-edge pairs. Specifically, every vertex stores its 3D location and a reference to any one attached edge; every edge stores references to its two vertices and any one attached face-edge pair; every face stores references to its two cells and one attached face-edge pair; and every cell stores a reference to any one attached face. Each face-edge pair stores references to one edge *E* and one face *F* adjacent to one another, along with a fixed number of adjacency relationships useful for topological traversals. Specifically, they store references (*spin*) to the two face-edge pairs reached by spinning *F* around *E* clockwise and counter-clockwise (see Fig. 8) and to the two face-edge pairs (*clock*) reached by moving around *F* in clockwise and counter-clockwise directions from *E* (see Fig. 8). The face-edge pair also stores a bit (*direction*) indicating whether the orientation of the vertices on the edge is clockwise or counter-clockwise with respect to the face within the pair. These simple, fixed-size structures make it possible to execute efficient topological traversals of space through cell, face, edge, and vertex adjacency relationships in a manner similar to the winged-edge[51] and facet-edge structures.[52]

We build the winged-pair data structure for a 3D polygonal model using a binary space partition (BSP),[53] a recursive binary split of 3D space into convex polyhedral regions (*cells*) separated by planes. To construct the BSP, we recursively split cells by the planes of the input polygons using the method described in Ref. 54. We start with a single BSP cell containing the entire 3D space and consider polygons of the input 3D model one-by-one. For each BSP cell split by a polygon *P*, the corresponding winged-pair cell is split along the plane supporting *P*, and the faces and edges on the

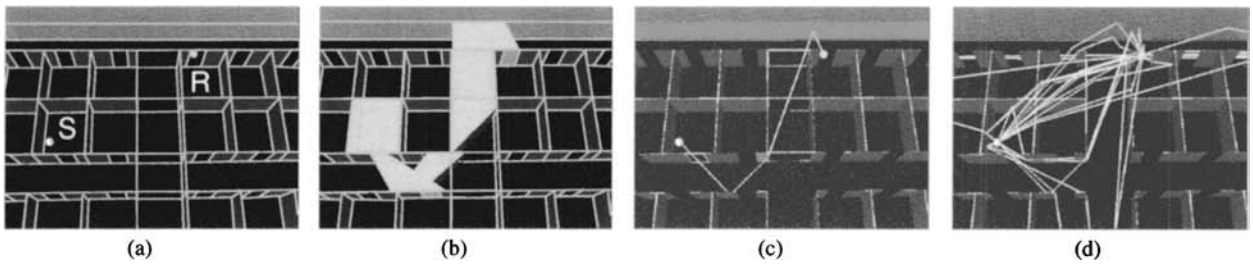


FIG. 7. Results of each phase of execution: (a) virtual environment (office cubicles) with source *S*, receiver *R*, and spatial subdivision marked in pink, (b) example reflected and diffracted beam (cyan) containing the receiver, (c) path generated for the corresponding sequence of opaque faces (green), transparent faces (purple), and edges (magenta), and (d) many paths found for different sequences from *S* to *R*.

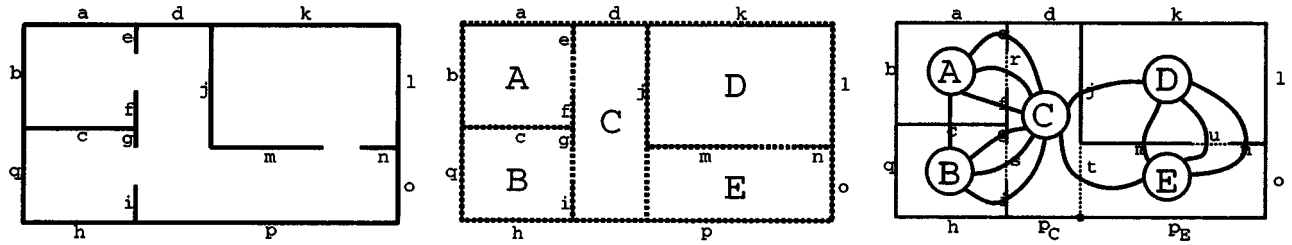


FIG. 9. Example spatial subdivision.

boundary of the cell are updated to maintain a three-manifold in which every face is flat, convex, and entirely inside or outside every input polygon. As faces are created, they are labeled according to whether they are reflectant (coincide with an input polygon) or transparent (split free space). The binary splitting process continues until no input polygon intersects the interior of any BSP cell, leading to a set of convex polyhedral cells whose faces are all convex and collectively contain all the input polygons. The resulting winged-pair is written to a file for use by later phases of our acoustic modeling process.

As an example, a simple 2D model (a) and its corresponding binary space partition (b) and cell adjacency graph (c) are shown in Fig. 9. Input ''polygons'' appear as solid line segments labeled with lower-case letters $(a-q)$; transparent cell boundaries introduced by the BSP are shown as dashed line segments labeled with lower-case letters $(r-u)$; constructed cell regions are labeled with upper-case letters $(A-E)$; and the cell adjacency graph implicit in the winged-pair structure is overlaid in Fig. 9(c).

B. Beam tracing

After the spatial subdivision has been constructed, we use it to accelerate traversals of space during beam tracing. The goal of this phase is to compute polyhedral beams representing the regions of space reachable from each stationary source by different sequences of reflections, transmissions, and diffractions. The beams are queried later during an interactive phase to compute propagation paths to specific receiver locations.

Briefly, beams are traced from each stationary sound source via a best-first traversal of the cell adjacency graph starting in the cell containing the source. As the algorithm traverses a cell boundary into a new cell, a copy of the current convex pyramidal beam is ''clipped'' to include only the region of space passing through the convex polygonal boundary to model transmissions. At each reflecting cell boundary, a copy of the transmission beam is mirrored across the plane supporting the cell boundary to model specular reflections. At each diffracting edge, a new beam is spawned whose source is the edge and whose extent includes all rays predicted by the geometric theory of diffraction.[55] The traversal along any sequence terminates when either the length of the shortest path within the beam or the cumulative attenuation exceed some user-specified thresholds. The traversal may also be terminated when the total number of beams traced or the elapsed time exceed other thresholds.

Pseudocode for the beam tracing algorithm appears in Fig. 10. Throughout the execution, a priority queue stores the set of beams to be traced sorted according to a *priority function*. Initially, the priority queue contains only one beam representing the entire space inside the cell containing the source. During each step of the algorithm, the highest priority beam B traversing a cell C is removed from the priority queue, and new ''child'' beams are placed onto the priority queue according to the following criteria:

```
void TraceBeams()
begin
    // Initialization
    S = Source point;
    D = Spatial subdivision;
    B = Beam containing all of space;
    C = Current cell;
    Q = Queue of beam tree nodes;

    C = FindCell(D, S);
    N = CreateNode(NULL, B, C);
    Q = InitQueue();

    PushQueue(Q, N);
    while (N = PopQueue(Q)) do
        // Consider each polygon on cell boundary
        foreach polygon P on boundary of N.C do
            // Check if polygon intersects beam
            if (Intersects(P, N.B)) then
                // Compute intersection beam
                Bt = Intersection(B, Beam(S, P));

                // Iterate along transmission paths
                if (Transmissive(P)) then
                    Ct = NeighborCell(D, C, P);
                    PushQueue(Q, CreateNode(N, Bt, Ct));
                endif

                // Iterate along reflection paths
                if (Reflective(P)) then
                    Br = Mirror(Bt, P);
                    PushQueue(Q, CreateNode(N, Br, C));
                endif

                // Iterate along diffraction paths
                foreach edge E on boundary of P do
                    // Check if edge intersects beam
                    if (Intersects(E, N.B)) then
                        Bd = CreateBeam(E);
                        PushQueue(Q, CreateNode(N, Bd, C));
                    endif
                endfor
            endif
        endfor
    endwhile
end
```

FIG. 10. Pseudocode for the beam tracing algorithm.

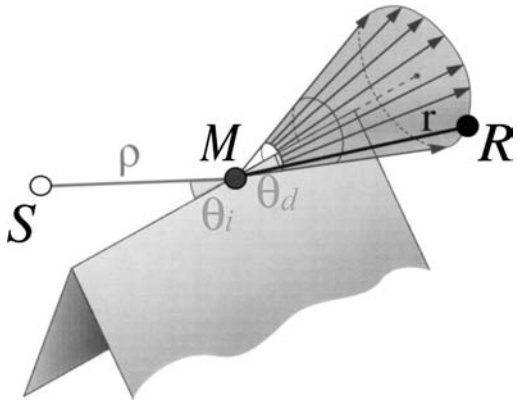


FIG. 11. Rays diffracted by a 3D edge according to the uniform theory of diffraction. The angle θ_d of the cone is equal to the angle θ_i between the incident ray and the edge.

(i) **Transmission beams**: For each nonopaque face F on the boundary of cell C and intersected by the beam B, a pyramidal *transmission beam* B_t is constructed whose apex is the source of B and whose sides each contain an edge of $F \cap B$. This new beam B_t represents sound rays traveling along B that are transmitted through F into the cell C_t which is adjacent to C across F.

(ii) **Specular reflection beams**: For each reflective face F on the boundary of cell C and intersected by the beam B, a polyhedral *specular reflection beam* B_r is constructed whose apex is the virtual source of B, created by mirroring the source of B over the plane containing F, and whose sides each contain an edge of $F \cap B$. This new beam B_r represents sound rays traveling along B that reflect specularly off of F and back into cell C.

(iii) **Diffraction beams**: For each edge E shared by two scattering faces F_1 and F_2 on the boundary of cell C and intersected by beam B, a *diffraction beam* is formed whose source is the line segment describing E and whose polyhedral extent contains the cone of potential diffraction paths bounded by the solid wedge of opaque surfaces sharing E, as shown in Fig. 11. This conservatively approximate beam contains all potential paths of sound initially traveling along B and then diffracted by edge E. For efficiency, the user may specify that diffraction beams should be traced only into shadow regions, in which case an extra half-space representing the shadow boundary is added to the beam.

Figure 12 contains an illustration of the beam tracing algorithm execution for the simple 2D example model shown in Fig. 9. The best-first traversal starts in the cell (labeled "D") containing the source point (labeled "S") with a beam containing the entire cell (D). Beams are created and traced for each of the six boundary polygons of cell "D" (j, k, l, m, n, and u). For example, transmission through the cell boundary labeled "u" results in a beam (labeled T_u) that is trimmed as it enters cell "E." T_u intersects only the polygon labeled "o," which spawns a reflection beam (labeled $T_u R_o$). That beam intersects only the polygon labeled "p," which spawns a reflection beam (labeled $T_u R_o R_p$), and so on.

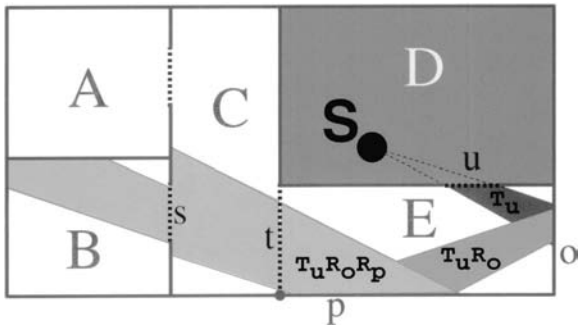


FIG. 12. Beam tracing through cell adjacency graph (this figure shows only one beam, while many such beams are traced along different propagation sequences).

Figure 13 shows an example in 3D with one sequence of beams traced up to one reflection from a source (on left) through the spatial subdivision (thin lines are cell boundaries) for a simple set of input polygons.

If the source is not a point, but instead distributed in a region of space (e.g., for diffracting edges), the exact region of space reachable by rays transmitted or reflected by a sequence of convex polygons can become quite complex, bounded by quadric surfaces corresponding to triple-edge (EEE) events.[56] Rather than representing these complex regions exactly, we conservatively overestimate the potential space of paths from each region of space edge with a convex polyhedron bounded by a fixed number of planes (usually six, as in Ref. 39). We correct for this approximation later during path generation by checking each propagation path to determine if it lies in the overestimating part of the polyhedron, in which case it is discarded. Since propagation patterns can be approximated conservatively and tightly with simple convex polyhedra, and since checking propagation paths is quick, the whole process is much more robust and faster than computing the exact propagation pattern directly. Using the adjacency information in the winged-pair structure, each new beam is constructed in constant time.

The results of the beam tracing algorithm are stored in a *beam tree* data structure[28] to be used later during path generation for rapid determination of propagation paths from the

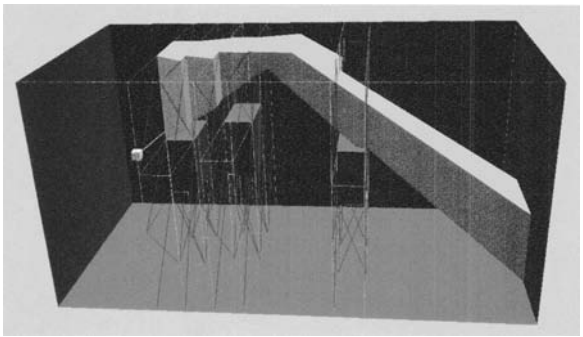


FIG. 13. A beam clipped and reflected at cell boundaries (this figure shows only one beam, while many such beams are traced along different propagation sequences).

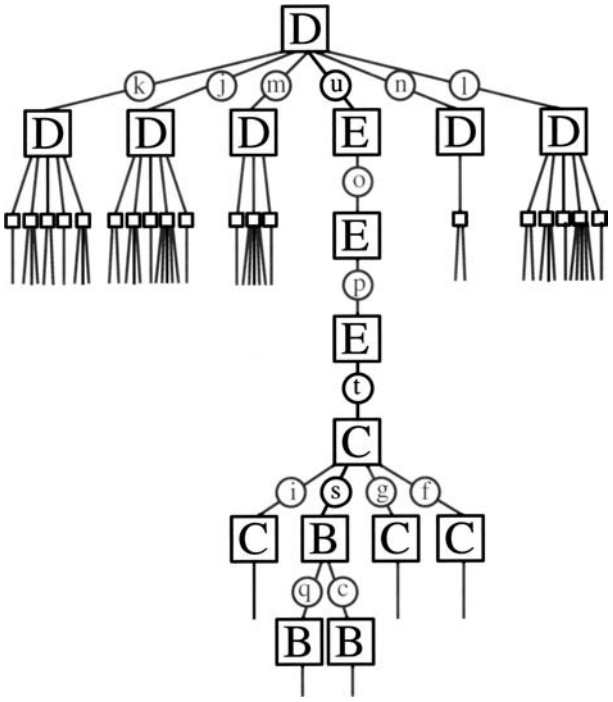


FIG. 14. Beam tree.

source point. The beam tree contains a node for each beam considered during the beam tracing algorithm. Specifically, each node stores (1) a reference to the cell being traversed, (2) a reference to the edge/face most recently traversed (if there is one), and (3) the convex polyhedral beam representing the region of space potentially reachable by the traversed sequence of transmissions, reflections, and diffractions. To further accelerate evaluation of propagation paths during a later interactive phase, each node of the beam tree also stores the cumulative attenuation due to reflective and transmissive absorption, and each cell of the spatial subdivision stores a list of ''back-pointers'' to its beam tree nodes. Figure 14 shows a partial beam tree corresponding to the traversal shown in Fig. 12.

C. Path generation

In the third phase, as a user moves the receiver interactively through the environment, we use the precomputed beam trees to identify propagation sequences of transmissions, reflections, and diffractions potentially reaching the receiver location.

Since every beam contains all points potentially reachable by rays traveling along a particular propagation sequence, we can quickly enumerate the potential propagation sequences by finding all the beams containing the receiver location. Specifically, we first find the cell containing the receiver by a logarithmic-time search of the BSP. Then, we check each beam tree node, T, associated with that cell to see whether the beam stored with T contains the receiver. If it does, a potential propagation sequence from the source point to the receiver point has been found, and the ancestors of T in the beam tree explicitly encode the set of reflections, diffractions, and transmissions through the boundaries of the

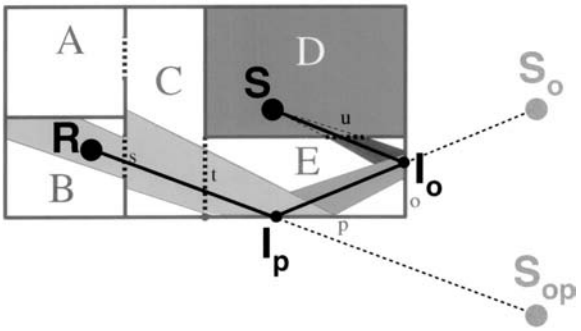


FIG. 15. Propagation path to receiver point (''R'') for the example in Figs. 12 and 14 computed via lookup in beam tree for source point (''S'').

spatial subdivision that a ray must traverse from the source to the receiver along this sequence (more generally, to any point inside the beam stored with T).

For each such propagation sequence, we construct explicit propagation path(s) from the source to the receiver. In our current system, we compute a single propagation path for each sequence as the one that is the shortest among all possible piecewise-linear paths from the source to the receiver (this path is used directly for modeling transmission, specular reflection, and diffraction according to the geometrical theory of diffraction). In order to construct this shortest path, we must find the points of intersection of the path with every face and edge in the sequence.

For sequences containing only transmissions and specular reflections (i.e., no edge diffractions), the shortest propagation path is generated analytically by iterative intersection with each reflecting surface. For instance, to find a path between a specific pair of points, S and R, along a sequence of specularly reflecting polygons P_i for $i=1,...,n$, we first traverse the polygon sequence in forward order to construct a stack of mirror images of S, where S_i corresponds to the image resulting from mirroring S over the first i of the n reflecting polygons in the sequence. Then, we construct the propagation path by traversing the polygon sequence in backward order, computing the ith vertex, V_i, of the path as the intersection of the line between V_{i-1} and S_{n-i+1} with the surface of polygon P_{n-i+1}, where V_0 is the receiver point. If every vertex V_i of the path lies within the boundary of the corresponding polygon P_i, we have found a *valid* reflection path from S to R along P. Otherwise, the path is in an overestimating part of the beam, and it can be ignored. Figure 15 shows the valid specular reflection path from the source (labeled ''S'') to a receiver (labeled ''R'') for the example shown in Fig. 12.

For sequences also containing diffracting edges, construction of the shortest propagation path is more difficult since it requires determining the locations of ''diffraction points,'' $D_i\,(i=1,...,n)$, for the n diffracting edges. These diffraction points generally lie in the interior of the diffracting edges (see Fig. 16), and the path through them locally satisfies a simple ''unfolding property:'' the angle (θ_i) at which the path enters each diffracting edge must be the same as the angle (ϕ_i) at which it leaves.[57] (The unfolding property is a consequence of the generalized Fermats principle.[55]) Thus, to find the shortest path through n diffracting edges

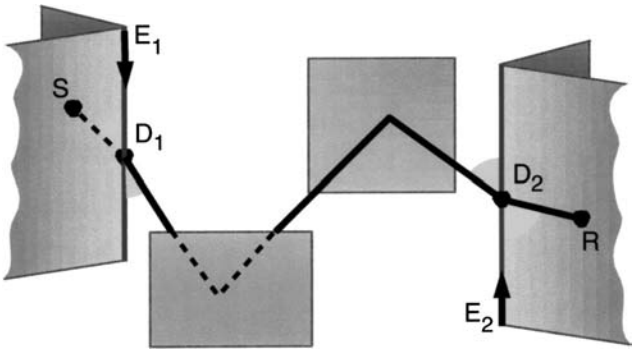


FIG. 16. A single propagation path comprising a diffraction, two specular reflections, and another diffraction. The two diffraction points (D_i) are determined by equal angle constraints at the corresponding edges (E_i).

and m transmitting and/or specularly reflecting faces, we must solve a nonlinear system of n equations expressing equal angle constraints at all diffracting edges:

$$\begin{aligned}
\overrightarrow{D_1S}\cdot\vec{E_1} &= \overrightarrow{D_1D_2}\cdot(-\vec{E_1})\\
\overrightarrow{D_2D_1}\cdot\vec{E_2} &= \overrightarrow{D_2D_3}\cdot(-\vec{E_2})\\
&\vdots\\
\overrightarrow{D_nD_{n-1}}\cdot\vec{E_n} &= \overrightarrow{D_nR}\cdot(-\vec{E_n})
\end{aligned} \tag{1}$$

where S is the source position, R is the receiver position, $\vec{E_i}$ is the normalized direction vector of the ith diffracting edge, and $\overrightarrow{D_{i+1}D_i}$ is a normalized direction vector between two adjacent points in the shortest path. To incorporate specular reflections in this equation, $\vec{E_i}$ and $\overrightarrow{D_{i+1}D_i}$ are both transformed by a mirroring operator accounting for the sequence of specularly reflecting faces up to the ith diffraction.

Parametrizing the edges, $D_i = O_i + t_i\vec{E_i}$ (where O_i is a reference point on edge i), the system of equations (1) can be rewritten in terms of n unknowns (t_i) and solved within a specified tolerance using a nonlinear system solving scheme. We use a locally convergent Newton scheme,[58] with the middle of the edges as a starting guess for the diffraction points. Since the equation satisfied by any diffraction point only depends on the previous and next diffraction points in the sequence, the Jacobian matrix is tridiagonal and can easily be evaluated analytically. Thus, every Newton iteration can be performed in time $O(n)$ where n is the number of unknowns (i.e., edges). We found this method to be faster than the recursive geometrical construction proposed by Aveneau.[59]

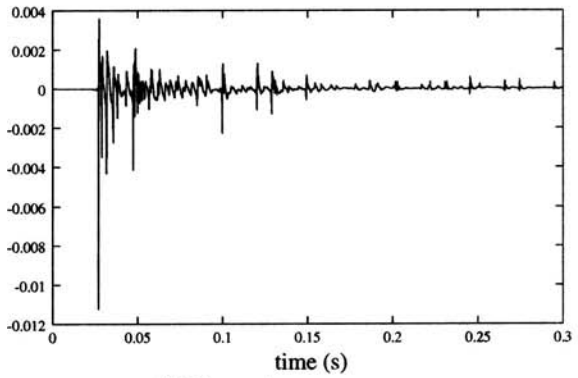
Once the intersection points of a propagation path are found, we validate whether the path intersects every surface and edge in the sequence (to compensate for the fact that the beams are conservatively approximate). If not, the path belongs to the overestimating part of the beam and is discarded. Otherwise, it contributes to an *impulse response* used for spatializing sound.

D. Auralization

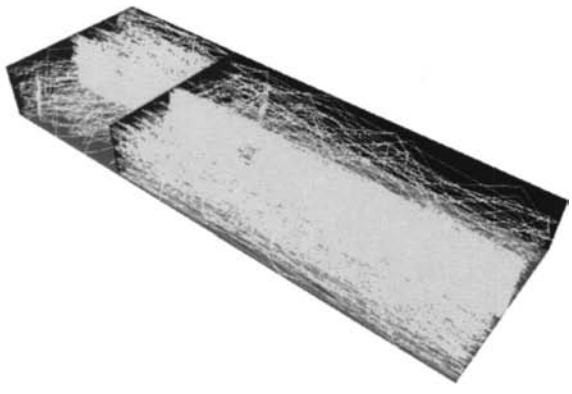
The computed early propagation paths are combined with a statistical approximation of late reverberation to model an impulse response of the virtual environment for every source/receiver pair. Dry sound signals emanating from the source are convolved with this digital filter to produce spatialized audio output.

Although this paper focuses on computation of geometric propagation paths, for the sake of completeness, we describe two auralization methods implemented in our system: (1) an off-line, high-resolution method for applications in which accuracy is favored over speed, and (2) an on-line, low-resolution approximation suitable for interactive walkthrough applications. Please refer to other papers (e.g., Refs. 7 and 60) for more details on auralization methods.

In the off-line case, we compute the early part of the impulse response in the Fourier frequency domain at the sampling rate resolution (e.g., 8000 complex values are updated for every propagation path for a one second long response at 16 kHz). As an example, Fig. 17 shows an impulse response computed for up to ten orders of specular reflections between source and receiver points in coupled-rooms. Our implementation includes frequency-dependent source and head filtering effects (obtained through measurements) and material filtering effects (derived from either measure-



(a) Impulse response



(b) Propagation paths

FIG. 17. Impulse response (left) computed for up to ten orders of specular reflections (right) between a point source (B&K "artificial mouth") and point receiver (omnidirectional) in a coupled-rooms environment (two rooms connected by an open door). There are 353 paths. The small room is $7\times8\times3$ m^3, while the large room is $17\times8\times3$ m^3.

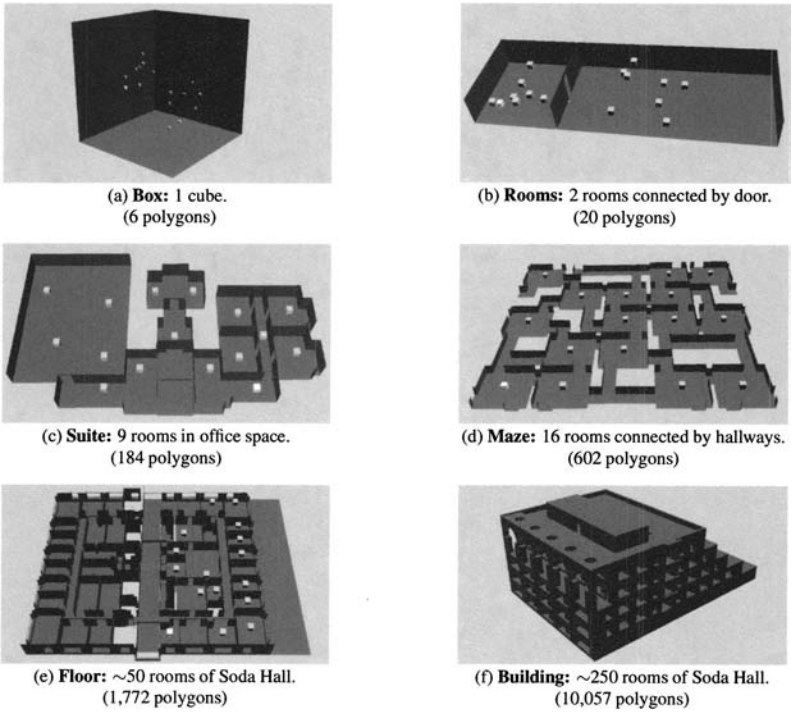


FIG. 18. Test models (source locations are gray dots).

ments or analytical models). We derive analytical models for the frequency-dependent impedance from the Delany–Bazley formula[61] and for the pressure reflection coefficient from the well-known plane wave formula of Pierce[62] (p. 33). Other wave formulas, such as the exact expression for the reflection of a spherical wave off an infinite impedant surface, derived by Thomasson,[63] could be used for improved accuracy in the near field from the surfaces.

We compute frequency-dependent diffraction coefficients using the uniform theory of diffraction[55,64,65] applied along the shortest paths constructed by our algorithm. If more accuracy is required, the information given in computed propagation sequences (exact intersected portions of surfaces and edges) can be used to derive filters, for example based on more recent results exploiting the Biot–Tolstoy–Medwin approach.[66–69] In this case, the shortest path computed by our algorithms can still be used to determine efficiently a unique incident direction on the listener's head for binaural processing (as suggested in Ref. 69).

In the on-line case, our system auralizes sound in real-time as the receiver position moves under interactive user control. A separate, concurrently executing process is spawned to perform convolution in software. To provide plausible spatialization with limited processing resources, we use a small number of frequency bands to reequalize and delay the source signal for every path, computing its contribution to a stereo impulse response in the time domain.[60] The delay associated with each path is given by L/C, where L is the length of the corresponding propagation path, and C is the speed of sound. The amplitude is given by A/L, where A is the product of all the attenuation coefficients for the reflecting, diffracting, and transmitting surfaces along the corresponding propagation sequence. Stereo impulse responses are generated by multiplying the amplitude of each path by the cardioid directivity function $((1+\cos(\theta))/2)$, where θ is the angle of arrival of the pulse with respect to the normal vector pointing out of the ear) corresponding to each ear. These gross approximations enable our auralization to give real-time feedback with purely software convolution. Other methods utilizing DSP hardware (e.g., binaural presentation) could easily be incorporated into our system in the future.

V. RESULTS

The 3D data structures and algorithms described in the preceding sections have been implemented in C++ and run on Silicon Graphics and PC/Windows computers.

To test whether the algorithms support large 3D environments and update propagation paths at interactive rates, we performed a series of experiments in which propagation paths were computed in a variety of architectural models (shown in Fig. 18). The test models ranged from a simple box with 6 polygons to a complex building with over 10 000 polygons. The experiments were run on a Silicon Graphics Octane workstation with 640 MB of memory and used one 195 MHz R10000 processor.

The focus of the experiments is to compare the computational efficiency of our method with image source methods, the approach most commonly used for interactive acoustic modeling applications. Accordingly, for the sake of direct comparison, we limited our beam tracing system to consider only specular reflections. In this case, our beam tracing method produces exactly the same set of propagation paths as classical image source methods. However, as we shall see, our beam tracing method has the ability to scale to large environments and to generate propagation paths at interactive rates.

In each experiment, we measured the time and storage required for spatial subdivision, beam tracing, sequence con-

TABLE I. Spatial subdivision statistics.

Model name	Input polys	Cell regions	Cell boundaries	Time (s)	Storage (MB)
Box	6	7	18	0.0	0.004
Rooms	20	12	43	0.1	0.029
Suite	184	98	581	3.0	0.352
Maze	602	172	1187	4.9	0.803
Floor	1772	814	5533	22.7	3.310
Bldg	10 057	4512	31 681	186.3	18.694

struction, and path generation. Results are reported in the following subsections.

A. Spatial subdivision results

We first constructed the spatial subdivision data structure (cell adjacency graph) for each test model. Statistics from this phase of the experiment are shown in Table I. Column 2 lists the number of input polygons in each model, while columns 3 and 4 contain the number of cell regions and boundary polygons, respectively, generated by the spatial subdivision algorithm. Column 5 contains the wall-clock execution time (in seconds) for the algorithm, while column 6 shows the storage requirements (in MBs) for the resulting spatial subdivision.

Empirically, we find that the number of cell regions and boundary polygons grows linearly with the number of input polygons for typical architectural models (see Fig. 19), rather than quadratically as is possible for worst case geometric arrangements. The reason for linear growth is illustrated in the two images inlaid in Fig. 19, which compare spatial subdivisions for the Maze test model (on the left) and a 2×2 grid of Maze test models (on the right). The 2×2 grid of Mazes has exactly four times as many polygons and approximately four times as many cells. The storage requirements of the spatial subdivision data structure also grow linearly as they are dominated by the vertices of boundary polygons.

The time required to construct the spatial subdivisions grows super-linearly, dominated by the code that selects and orders splitting planes during BSP construction (see Ref. 54). However, it is important to note that the spatial subdivision phase need be executed only once off-line for each geometric model, as its results are stored in a file, allowing rapid reconstruction in subsequent beam tracing executions.

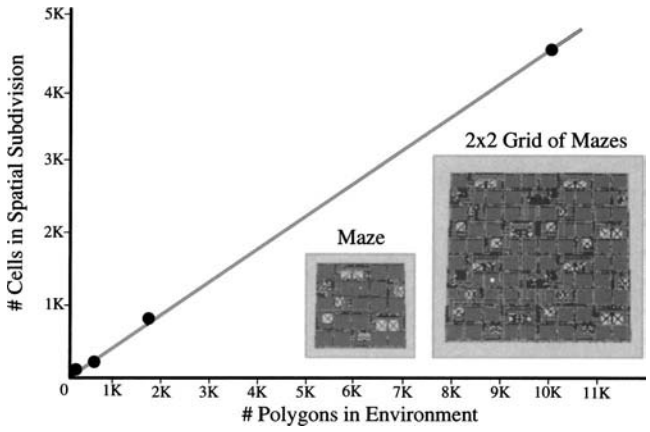


FIG. 19. Plot of subdivision size versus polygonal complexity.

TABLE II. Beam tracing and path generation statistics.

Model name	No. of polys	No. of Rfl	Beam tracing		Path generation	
			No. of beams	Time (ms)	No. of paths	Time (ms)
Box	6	0	1	0	1.0	0.0
		1	7	1	7.0	0.1
		2	37	3	25.0	0.3
		4	473	42	129.0	6.0
		8	10 036	825	833.0	228.2
Rooms	20	0	3	0	1.0	0.0
		1	31	3	7.0	0.1
		2	177	16	25.1	0.3
		4	1939	178	127.9	5.2
		8	33 877	3024	794.4	180.3
Suite	184	0	7	1	1.0	0.0
		1	90	9	6.8	0.1
		2	576	59	25.3	0.4
		4	7217	722	120.2	6.5
		8	132 920	13 070	672.5	188.9
Maze	602	0	11	1	0.4	0.0
		1	167	16	2.3	0.0
		2	1162	107	8.6	0.1
		4	13 874	1272	36.2	2.0
		8	236 891	21 519	183.1	46.7
Floor	1772	0	23	4	1.0	0.0
		1	289	39	6.1	0.1
		2	1713	213	21.5	0.4
		4	18 239	2097	93.7	5.3
		8	294 635	32 061	467.0	124.5
Bldg	10 057	0	28	5	1.0	0.0
		1	347	49	6.3	0.1
		2	2135	293	22.7	0.4
		4	23 264	2830	101.8	6.8
		8	411 640	48 650	529.8	169.5

B. Beam tracing results

We tested our beam tracing algorithm with 16 source locations in each test model. The source locations were chosen to represent typical audio source positions (e.g., in offices, in common areas, etc.)—they are shown as gray dots in Fig. 18 (we use the same source locations in Building model as in the Floor model). For each source location, we traced beams (i.e., constructed a beam tree) five times, each time with a different limit on the maximum order of specular reflections (e.g., up to 0, 1, 2, 4, or 8 orders). Other termination criteria based on attenuation or path length were disabled, and transmission was ignored, in order to isolate the impact of input model size and maximum order of specular reflections on computational complexity.

Table II contains statistics from the beam tracing experiment—each row represents a test with a particular 3D model and maximum order of reflections, averaged over all 16 source locations. Columns 2 and 3 show the number of polygons describing each test model and the maximum order of specular reflections allowed in each test, respectively. Column 4 contains the average number of beams traced by our

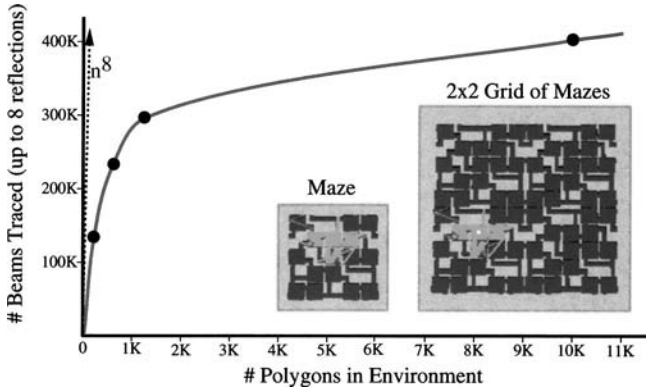


FIG. 20. Plot of beam tree size versus polygonal complexity.

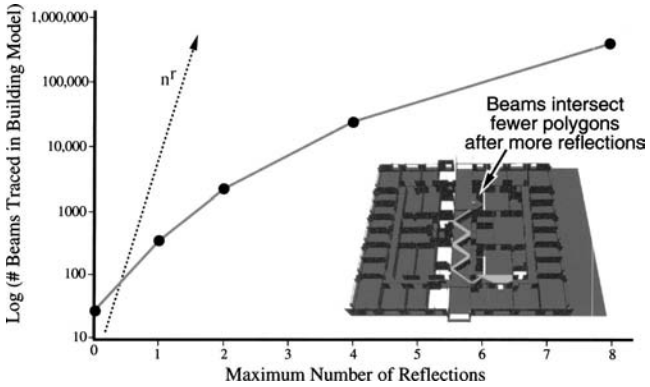


FIG. 21. Plot of beam tree size with increasing reflection orders.

algorithm (i.e., the average number of nodes in the resulting beam trees), and column 5 shows the average wall-clock time (in milliseconds) for the beam tracing algorithm to execute.

1. Scale with increasing polygonal complexity

We readily see from the results in column 4 that the number of beams traced by our algorithm (i.e., the number of nodes in the beam tree) does *not* grow at an exponential rate with the number of polygons (n) in these environments (as it does using the image source method). Each beam traced by our algorithm preclassifies the regions of space according to whether the corresponding virtual source (i.e., the apex of the beam) is visible to a receiver. Rather than generating a virtual source (beam) for every front-facing surface at each step of the recursion as in the image source method, we directly find only the potentially visible virtual sources via beam-polygon intersection and cell adjacency graph traversal. We use the current beam and the current cell of the spatial subdivision to find the small set of polygon reflections that admit visible higher-order virtual sources.

The benefit of this approach is particularly important for large environments in which the boundary of each convex cell is simple, and yet the entire environment is very complex. As an example, consider computation of up to eighth order specular reflections in the Building test model (the last row of Table II). The image source method must consider approximately 1 851 082 741 virtual sources ($\Sigma_{r=0}^{8}(10\,057/2)^r$), assuming half of the 10 057 polygons are front-facing to each virtual source. Our beam tracing method considers only 411 640 virtual sources, a difference of four orders of magnitude. In most cases, it would be impractical to build and store the recursion tree without such effective pruning.

In ''densely occluded'' environments, in which all but a little part of the environment is occluded from any source point (e.g., most buildings and cities), the number of beams traced by our algorithm even grows sublinearly with the total number of polygons in the environment (see Fig. 20). In these environments, the number of sides to each polyhedral cell is nearly constant, and a nearly constant number of cells are reached by each beam, leading to near-constant expected-case complexity of our beam tracing algorithm with increasing global environment complexity.

This result is most readily understood by comparing the number of beams traced for up to eighth order reflections in the Floor and Building models (i.e., the rightmost two data points in Fig. 20). The Floor model represents the fifth floor of Soda Hall at UC Berkeley. It is a subset of the Building model, which represents five floors of the same building (floors 3–7 and the roof). Although the Building model (10 057 polygons) has more than five times the complexity of the Floor model (1772 polygons), the average number of beams traced from the same source locations by our algorithm is only 1.2–1.4 times larger for the Building model (e.g., 411 640/294 635=1.4). This is because the complexity of the spatial subdivision on the fifth floor of the building is similar in both cases, and most other parts of the building are not reached by any beam. Similarly, we expect that the beam tracing algorithm would have nearly the same complexity if the entire building were 1000 floors high, or if it were in a city of 1000 buildings. This result is shown visually in Fig. 20: the number of beams (green) traced in the Maze test model (left) does not increase significantly if the model is increased to be a 2×2 grid of Maze models (right). *The beam tracing algorithm is impacted only by local complexity, and not by global complexity.*

2. Scale with increasing reflections

We see that the number of beams traced by our algorithm grows exponentially as we increase the maximum order of reflections (r), but far slower than $O(n^r)$ as in the image source method. Figure 21 shows a logscale plot of the average number of beams traced in the Building model with increasing orders of specular reflections. The beam tree growth is less than $O(n^r)$ because each beam narrows as it is clipped by the cell boundaries it has traversed, and thus it tends to intersect fewer cell boundaries (see the example beam inlaid in Fig. 21). In the limit, each beam becomes so narrow that it intersects only one or two cell boundaries, on average, leading to a beam tree with a small branching factor (rather than a branching factor of $O(n)$, as in the image source method).

As an example, consider Table III which shows the average branching factor for nodes at each depth of the beam tree constructed for up to eighth order specular reflections in the Building model from one source location. The average

TABLE III. Example beam tree branching statistics. The number of interior nodes (have children) and leaf nodes (have no children) at each depth are listed, along with the average branching factor for interior nodes.

Tree depth	Total nodes	Interior nodes	Leaf nodes	Branching factor
0	1	1	0	16.0000
1	16	16	0	6.5000
2	104	104	0	4.2981
3	447	446	1	2.9193
4	1302	1296	6	2.3920
5	3100	3092	8	2.0715
6–10	84 788	72 469	12 319	1.2920
11–15	154 790	114 664	40 126	1.2685
>15	96 434	61 079	35 355	1.1789

branching factor (column 5) generally decreases with tree depth and is generally bounded by a small constant in lower levels of the tree.

C. Path generation results

In order to verify that specular reflection paths are computed at interactive rates from stationary sources as the receiver moves, we conducted experiments to quantify the complexity of generating specular reflection paths to different receiver locations from precomputed beam trees. For each beam tree in the previous experiment, we logged statistics during generation of specular propagation paths to 16 different receiver locations. Receivers were chosen randomly within a two foot sphere around the source to represent a typical audio scenario in which the source and receiver are in close proximity within the same ''room.'' We believe this represents a worst-case scenario as fewer paths would likely reach more remote and more occluded receiver locations.

Columns 6 and 7 of Table II contain statistics gathered during path generation for each combination of model and termination criterion averaged over all 256 source-receiver pairs (i.e., 16 receivers for each of the 16 sources). Column 6 contains the average number of propagation paths generated, while column 7 shows the average wall-clock time (in milliseconds) for execution of the path generation algorithm. Figure 22 shows a plot of the wall-clock time required to generate up to eighth order specular reflection paths for each test model.

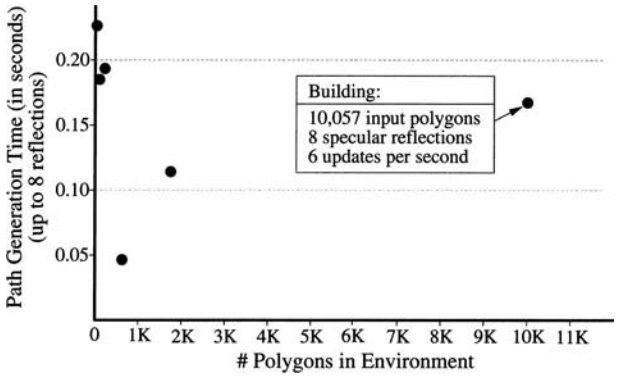


FIG. 22. Path compute time versus polygonal complexity.

We find that the number of specular reflection paths is nearly constant across all of our test models when a source and receiver are located in close proximity of one another. Also, the time required by our path generation algorithm is generally *not* dependent on the number of polygons in the environment (see Fig. 22), nor is it dependent on the total number of nodes in the precomputed beam tree. This result is due to the fact that our path generation algorithm considers only nodes of the beam tree with beams residing inside the cell containing the receiver location. Therefore, the computation time required by the algorithm is *not* dependent on the complexity of the environment outside the receiver's cell, but instead on the number of beams that traverse the receiver's cell.

Overall, we find that our algorithm supports generation of specular reflection paths between a fixed source and any (arbitrarily moving) receiver at interactive rates in complex environments. For instance, we are able to compute up to eighth order specular reflection paths in the Building environment with more than 10 000 polygons at a rate of approximately six times per second (i.e., the rightmost point in the plot of Fig. 22).

VI. DISCUSSION

In this paper, we describe beam tracing algorithms and data structures that accelerate computation of propagation paths in large architectural environments. The following subsections discuss applications of the proposed methods, limitations of our approach, and related topics for further study.

A. Applications

There are several potential applications for the methods proposed in this paper. For instance, traditional accoustical design programs (e.g., CATT Acoustics[12]) could be enhanced with real-time auralization and visualization that aid a user in understanding which surfaces cause particular acoustical effects.

Alternatively, real-time acoustic simulation can be used to enhance simulation of virtual environments in interactive walkthrough applications. Auditory cues are important in immersive applications as they can combine with visual cues to

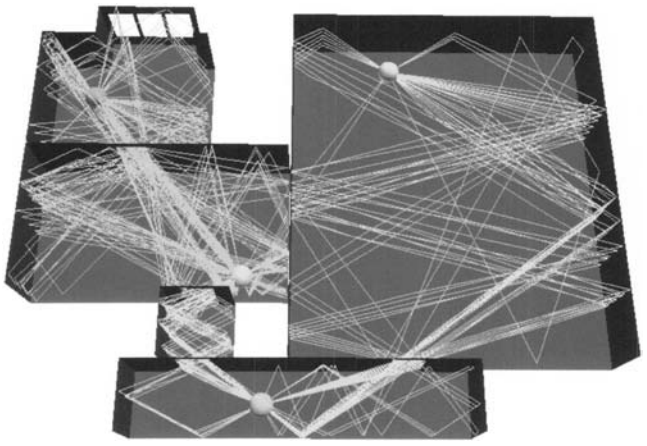


FIG. 23. Sound propagation paths (lines) between four avatars (spheres) representing users in shared virtual environment.

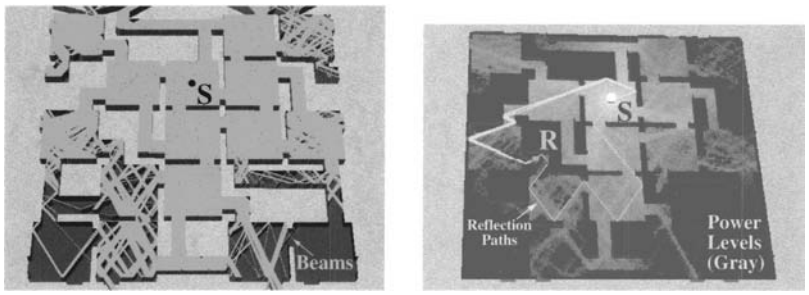


FIG. 24. Eighth-order specular reflection beams (left) and predicted power levels (right) in Maze model.

aid localization of objects, separation of simultaneous sound signals, and formation of spatial impressions of an environment.[70] For instance, binaural auditory cues are helpful in localizing objects outside a user's field of view, such as when a car comes around a blind corner in a driving simulation. They also help the separation of simultaneous sounds (e.g., many speakers at a cocktail party). Finally, qualitative changes in sound propagation, such as more absorption in a room with more plush carpets, can enhance and reinforce visual comprehension of the environment. Experiments have shown that more accurate acoustic modeling provides a user with a stronger sense of presence in a virtual environment.[2]

We have integrated our beam tracing method into an immersive system that allows a user to move through a virtual environment while images and spatialized audio are rendered in real-time according to the user's simulated viewpoint.[8–10] In the example shown in Fig. 23, multiple users represented by avatars (spheres) sharing a virtual world can speak to one another while the system spatializes their voices according to sound propagation paths (lines) through the environment.

In order to support multiple simultaneously moving sources and receivers,[9] as is required by a distributed virtual environment application with many avatars, we can no longer precompute beam trees. Instead, we must compute them in real-time as the source moves. However, we can take advantage of the fact that sounds can only be generated or heard at the positions of ''avatars'' representing the users. This simple observation enables two important enhancements to our beam tracing method. First, a bidirectional beam tracing algorithm combines beams traced from both sources and receivers to find propagation paths between them. Second, an amortized beam tracing algorithm computes beams emanating from box-shaped regions of space containing predicted avatar locations and reuses those beams multiple times to compute propagation paths as each avatar moves inside the box. We have incorporated these two enhancements into a time-critical multiprocessing system that allocates its computational resources dynamically in order to compute the highest priority propagation paths between moving avatar locations in real-time with graceful degradation and adaptive refinement. These enhancements result in two orders of magnitude of improvement in computational efficiency in the case where beams are traced for known receiver locations. See Ref. 9 for details.

Overall, we find that precomputing beams is advantageous for stationary sound sources and arbitrarily moving receivers, while computing them asynchronously on the fly is still practical for continuously moving sources and receivers.

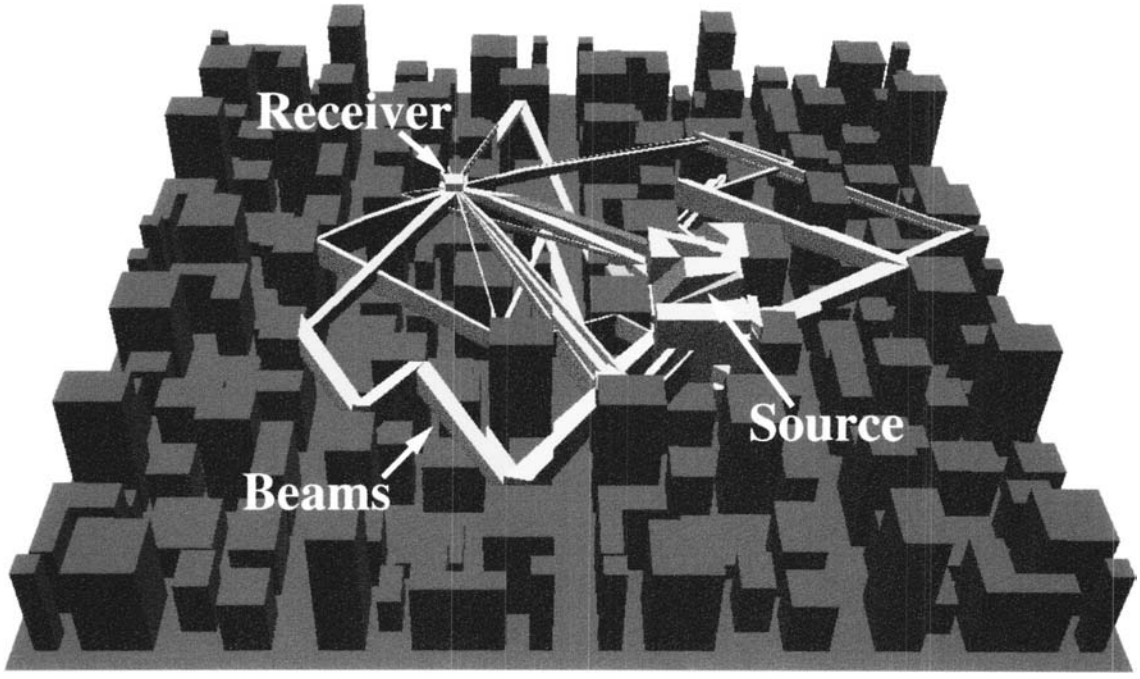


FIG. 25. Beams (green) containing all eighth-order specular reflection paths from a source to a receiver in City model.

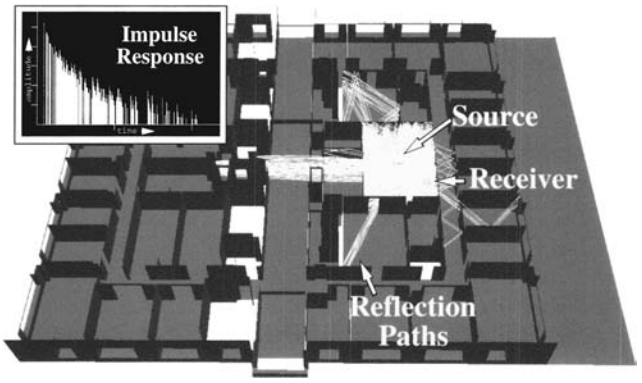


FIG. 26. Impulse response (inset) derived from eighth-order specular reflection paths (yellow) in Floor model.

B. Visualization

In order to aid the understanding and debugging of our acoustic modeling method, we find it extremely valuable to use interactive visualization of our data structures and algorithms. Our system provides menu and keyboard commands that may be used to toggle display of the (1) input polygons, (2) source point, (3) receiver point, (4) boundaries of the spatial subdivision, (5) pyramidal beams, (6) image sources, and (7) propagation paths. The system also supports visualization of acoustic metrics (e.g., power, clarity, etc.) for a set of receiver locations on a regular planar grid displayed with a textured polygon. Example visualizations are shown in Figs. 24–26.

Of course, many commercial[11–13] and research systems[29,71] provide elaborate tools for visualizing computed acoustic metrics. The critical difference in our system is that it supports continuous interactive updates of propagation paths and debugging information as a user moves the receiver point with the mouse. For instance, Figs. 24 and 26 show eighth-order specular reflection paths from a single audio source to a receiver location which is updated more than six times per second as the receiver location is moved arbitrarily. Figure 27 shows paths with specular reflections and diffractions computed in a city model and an auditorium. The user may select any propagation path for further inspection by clicking on it and then independently toggle display of reflecting cell boundaries, transmitting cell boundaries, and the polyhedral beams associated with the selected path.

Separate pop-up windows provide real-time display of other useful visual debugging and acoustic modeling information. For instance, one window shows a diagram of the beam tree data structure. Each beam tree node is dynamically colored in the diagram according to whether the receiver point is inside its associated beam or cell. Another window shows a plot of the impulse response representing the propagation paths from source to receiver (see Fig. 26). A third window shows values of various acoustic metrics, including power, clarity, reverberation time, and frequency response. All of the information displayed is updated in real-time as the user moves the receiver interactively with the mouse.

C. Geometric limitations

Our system is a research prototype, and it has several limitations. First, the 3D model must comprise only planar polygons because we do not model the transformations for beams as they reflect off curved surfaces. Furthermore, we do not trace beams along paths of refraction or diffuse reflection, which may be important acoustical effects. Each acoustic reflector is assumed to be locally reacting and to have dimensions far exceeding the wavelength of audible sound.

Second, our methods are only practical for coarse 3D models without highly faceted surfaces, such as the ones often found in acoustic modeling simulations of architectural spaces and concert halls. The difficulty is that beams are fragmented by cell boundaries as they are traced through a cell adjacency graph. For this reason, our beam tracing method would not perform well for geometric models with high local geometric complexity (e.g., a forest of trees).

Third, the major occluding and reflecting surfaces of the virtual environment must be static through the entire execution. If any acoustically significant polygon were to move, the cell adjacency graph would have to be updated incrementally.

The class of geometric models for which our method does work well includes most architectural and urban environments. In these cases, acoustically significant surfaces are generally planar, large, and stationary, and the acoustical effects of any sound source are limited to a local region of the environment (*densely occluded*).

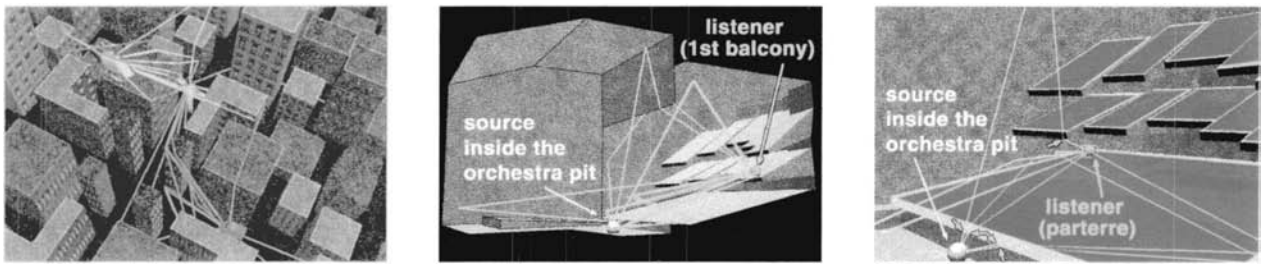


FIG. 27. Visualizations of sound paths in different environments. Diffraction of sound in a city environment is shown on the left, while early propagation paths for a source located in the orchestra pit of an opera house are shown in the middle and right images. Note the diffracted paths over the lip of the pit and balconies (cyan arrows).

D. Future work

Our system could be extended in many ways. For instance, the beam tracing algorithm is well suited for parallelization, with much of the previous work in parallel ray tracing directly applicable.[72] Also, the geometric regions covered by each node of the beam tree could be stored in a single hierarchical spatial structure (e.g., a BSP), allowing logarithmic search during path generation, rather than linear search of the beams inside a single cell. Of course, we could also use beam trees to allow a user to manipulate the acoustic properties of individual surfaces of the environment interactively with real-time feedback, such as for parametrized ray tracing[73] or for inverse modeling.[74]

Verification of our simulation results by comparison to measured data is an important topic for further study. In this paper, we have purposely focused on the computational aspects of geometrical acoustic modeling and left the validation of the models for future work. For this aspect of the problem, we refer the reader to related verification studies (e.g., Ref. 75), noting that our current system can compute the same specular reflection paths as methods based on image sources and ray tracing for which verification results are published (e.g., Refs. 76 and 77).

We are currently making impulse response measurements for verification of our simulations with reflections, diffractions, and transmissions. We have recently built a ''room'' for validation experiments. The base configuration of the room is a simple box. However, it is constructed with reconfigurable panels that can be removed or inserted to create a variety of interesting geometries, including ones with diffracting panels in the room's interior. We are currently measuring the directional reflectance distribution of each of these panels in the Anechoic Chamber at Bell Laboratories. We plan to make measurements with speakers and microphones at several locations in the room and with different geometric arrangements of panels, and we will compare the measurements with the results of simulations with the proposed beam tracing algorithms for matching configurations.

Perhaps the most interesting direction of future work is to investigate the possible applications of *interactive* acoustic modeling. What can we do with interactive manipulation of acoustic model parameters that would be difficult to do otherwise? As a first application, we hope to build a system that uses our interactive acoustic simulations to investigate the psychoacoustic effects of varying different acoustic modeling parameters. Our system will allow a user to interactively change various acoustic parameters with real-time auralization and visualization feedback. With this interactive simulation system, it may be possible to address psychoacoustic questions, such as ''how many reflections are psychoacoustically important to model?'' or ''which surface reflection model provides a psychoacoustically better approximation?'' Moreover, we hope to investigate the interaction of visual and aural cues on spatial perception. We believe that the answers to such questions are of critical importance to future design of 3D simulation systems.

VII. CONCLUSION

We have described a system that uses beam tracing data structures and algorithms to compute early propagation paths from static sources to a moving receiver at interactive rates for real-time auralization in large architectural environments.

As compared to previous acoustic modeling approaches, our beam tracing method takes unique advantage of *precomputation* and *convexity*. Precomputation is used twice, once to encode in the spatial subdivision data structure a depth-ordered sequence of (cell boundary) polygons to be considered during any traversal of space, and once to encode in the beam tree data structure the region of space reachable from a static source by sequences of specular reflections, diffractions, and transmissions at cell boundaries. We use the convexity of the beams, cell regions, and cell boundary polygons to enable efficient and robust computation of beam-polygon and beam-receiver intersections. As a result, our method is uniquely able to (1) enumerate all propagation paths robustly, (2) scale to compute propagation paths in large, densely occluded environments, (3) model effects of edge diffraction in arbitrary polyhedral environments, and (4) support evaluation of propagation paths at interactive rates. Our interactive system integrates real-time auralization with visualization of large virtual environments.

Based on our initial experiences with this system, we believe that interactive geometric acoustic modeling provides a valuable new tool for understanding sound propagation in complex 3D environments. We are continuing this research in order to further investigate the perceptual interaction of visual and acoustical effects and to better realize the opportunities possible with interactive acoustic modeling.

ACKNOWLEDGMENTS

The authors thank Sid Ahuja for his support of this work, and Arun C. Surendran and Michael Gatlin for their valuable discussions and contributions to the project.

[1] D. R. Begault, *3D Sound for Virtual Reality and Multimedia* (Academic, New York, 1994).
[2] N. I. Durlach and A. S. Mavor, *Virtual Reality Scientific and Technological Challenges*, National Research Council Report (National Academy, Washington, D.C., 1995).
[3] H. Kuttruff, *Room Acoustics (3rd edition)* (Elsevier Applied Science, New York, 1991).
[4] J. B. Allen and D. A. Berkley, ''Image method for efficiently simulating small room acoustics,'' J. Acoust. Soc. Am. **65**, 943–950 (1979).
[5] J. Borish, ''Extension of the image model to arbitrary polyhedra,'' J. Acoust. Soc. Am. **75**, 1827–1836 (1984).
[6] U. R. Krockstadt, ''Calculating the acoustical room response by the use of a ray tracing technique,'' J. Sound Vib. **8**(18), 118–125 (1968).
[7] M. Kleiner, B. I. Dalenback, and P. Svensson, ''Auralization—An overview,'' J. Audio Eng. Soc. **41**(11), 861–875 (1993).
[8] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, M. Sondhi, and J. West, ''A beam tracing approach to acoustic modeling for interactive virtual environments,'' *ACM Computer Graphics, SIGGRAPH'98 Proceedings*, July 1998, pp. 21–32.
[9] T. Funkhouser, P. Min, and I. Carlbom, ''Real-time acoustic modeling for distributed virtual environments,'' *ACM Computer Graphics, SIGGRAPH'99 Proceedings*, August 1999, pp. 365–374.
[10] N. Tsingos, T. Funkhouser, A. Ngan, and I. Carlbom, ''Modeling acoustics in virtual environments using the uniform theory of diffraction,'' *ACM Computer Graphics, SIGGRAPH 2001 Proceedings*, August 2001, pp. 545–552.

[11] Bose Corporation, Bose Modeler, Framingham, MA, http://www.bose.com.
[12] CATT-Acoustic, Gothenburg, Sweden, http://www.netg.se/catt.
[13] J. M. Naylor, ''Odeon—Another hybrid room acoustical model,'' Appl. Acoust. **38**(1), 131–143 (1993).
[14] R. D. Ciskowski and C. A. Brebbia (eds.), *Boundary Element Methods in Acoustics* (Elseiver Applied Science, New York, 1991).
[15] P. Filippi, D. Habault, J. P. Lefevre, and A. Bergassoli, *Acoustics, Basic Physics, Theory and Methods* (Academic, New York, 1999).
[16] A. Kludszuweit, ''Time iterative boundary element method (TIBEM)—a new numerical method of four-dimensional system analysis for the calculation of the spatial impulse response,'' Acustica **75**, 17–27 (1991) (in German).
[17] S. Kopuz and N. Lalor, ''Analysis of interior acoustic fields using the finite element method and the boundary element method,'' Appl. Acoust. **45**, 193–210 (1995).
[18] G. R. Moore, ''An Approach to the Analysis of Sound in Auditoria,'' Ph.D. thesis, Cambridge, UK, 1984.
[19] N. Tsingos and J.-D. Gascuel, ''Soundtracks for computer animation: Sound rendering in dynamic environments with occlusions,'' *Graphics Interface'97*, May 1997, pp. 9–16.
[20] M. F. Cohen and J. R. Wallace, *Radiosity and Realistic Image Synthesis* (Academic, New York, 1993).
[21] F. X. Sillion and C. Puech, *Radiosity and Global Illumination* (Morgan Kaufmann, San Francisco, 1994).
[22] K. H. Kuttruff, ''Auralization of impulse responses modeled on the basis of ray-tracing results,'' J. Audio Eng. Soc. **41**(11), 876–880 (1993).
[23] U. R. Kristiansen, A. Krokstad, and T. Follestad, ''Extending the image method to higher-order reflections,'' J. Appl. Acoust. **38**(2-4), 195–206 (1993).
[24] R. Cook, T. Porter, and L. Carpenter, ''Distributed ray-tracing,'' *ACM Computer Graphics, SIGGRAPH'84 Proceedings*, July 1984, Vol. 18(3), pp. 137–146.
[25] J. T. Kajiya, ''The rendering equation,'' *ACM Computer Graphics, SIGGRAPH'86 Proceedings*, Vol. 20(4), pp. 143–150.
[26] H. Lehnert, ''Systematic errors of the ray-tracing algorithm,'' Appl. Acoust. **38**, 207–221 (1993).
[27] N. Dadoun, D. G. Kirkpatrick, and J. P. Walsh, ''The geometry of beam tracing,'' *Proceedings of the Symposium on Computational Geometry*, June 1985, pp. 55–71.
[28] P. Heckbert and P. Hanrahan, ''Beam tracing polygonal objects,'' *ACM Computer Graphics, SIGGRAPH'84 Proceedings*, Vol. 18(3), pp. 119–127.
[29] M. Monks, B. M. Oh, and J. Dorsey, ''Acoustic simulation and visualisation using a new unified beam tracing and image source approach,'' *Proc. Audio Engineering Society Convention*, 1996, pp. 153–174.
[30] U. Stephenson and U. Kristiansen, ''Pyramidal beam tracing and time dependent radiosity,'' *Fifteenth International Congress on Acoustics*, June 1995, pp. 657–660.
[31] J. P. Walsh and N. Dadoun, ''What are we waiting for? The development of Godot, II,'' 103rd Meeting of the Acoustical Society of America, April 1982.
[32] J. H. Chuang and S. A. Cheng, ''Computing caustic effects by backward beam tracing,'' Visual Comput. **11**(3), 156–166 (1995).
[33] A. Fujomoto, ''Turbo beam tracing—A physically accurate lighting simulation environment,'' *Knowledge Based Image Computing Systems*, May 1988, pp. 1–5.
[34] G. Ghazanfarpour and J. M. Hasenfratz, ''A beam tracing with precise antialiasing for polyhedral scenes,'' Comput. Graph. **22**(1), 103–115 (1998).
[35] E. Haines, ''Beams O' Light: Confessions of a hacker,'' *Frontiers in Rendering, Course Notes, SIGGRAPH'91*, 1991.
[36] M. Watt, ''Light-water interaction using backward beam tracing,'' *ACM Computer Graphics, SIGGRAPH'90 Proceedings*, August 1990, pp. 377–385.
[37] C. B. Jones, ''A new approach to the 'hidden line' problem,'' Comput. J. **14**(3), 232–237 (1971).
[38] T. Funkhouser, ''A visibility algorithm for hybrid geometry- and image-based modeling and rendering,'' Comput. Graph. **23**(5), 719–728 (1999).
[39] S. Teller, ''Visibility Computations in Densly Occuded Polyhedral Environments,'' Ph.D. thesis, Computer Science Div., University of California, Berkeley, 1992.
[40] S. Fortune, ''Algorithms for prediction of indoor radio propagation,'' Technical Report Document 11274-960117-03TM, Bell Laboratories, 1996.
[41] S. J. Fortune, ''Topological beam tracing,'' in *Proc. 15th ACM Symposium on Computational Geometry*, 1999, pp. 59–68.
[42] J. Amanatides, ''Ray tracing with cones,'' *ACM Computer Graphics, SIGGRAPH'84 Proceedings*, July 1984, Vol. 18(3), pp. 129–135.
[43] J. P. Vian and D. van Maercke, ''Calculation of the room response using a ray tracing method,'' *Proceedings of the ICA Symposium on Acoustics and Theater Planning for the Performing Arts*, 1986, pp. 74–78.
[44] T. Lewers, ''A combined beam tracing and radiant exchange computer model of room acoustics,'' Appl. Acoust. **38**, 161–178 (1993).
[45] P. Kreuzgruber, P. Unterberger, and R. Gahleitner, ''A ray splitting model for indoor radio propagation associated with complex geometries,'' in *Proceedings of the 1993 43rd IEEE Vehicular Technology Conference*, 1993, pp. 227–230.
[46] A. Rajkumar, B. F. Naylor, F. Feisullin, and L. Rogers, ''Predicting RF coverage in large environments using ray-beam tracing and partitioning tree represented geometry,'' Wireless Networks **2**(2), 143–154 (1996).
[47] S. Teller and P. Hanrahan, ''Global visibility algorithms for illumination computations,'' pp. 239–246 (1993).
[48] S. Teller, C. Fowler, T. Funkhouser, and P. Hanrahan, ''Partitioning and ordering large radiosity computations,'' *ACM Computer Graphics, SIGGRAPH'93 Proceedings*, August 1994, pp. 443–450.
[49] P. Min and T. Funkhouser, ''Priority-driven acoustic modeling for virtual environments,'' *EUROGRAPHICS 2000*, August 2000, pp. 179–188.
[50] F. Durand, G. Drettakis, and C. Puech, ''The visibility skeleton: A powerful and efficient multi-purpose global visibility tool,'' in *Computer Graphics, SIGGRAPH'97 Proceedings*, 1997, pp. 89–100.
[51] B. G. Baumgart, ''Winged edge polyhedron representation,'' Technical Report AIM-179 (CS-TR-74-320), Computer Science Department, Stanford University, Palo Alto, CA, October 1972.
[52] D. P. Dobkin and M. J. Laszlo, ''Primitives for the manipulation of three-dimensional subdivisions,'' Algorithmica **4**(1), 3–32 (1989).
[53] H. Fuchs, Z. M. Kedem, and B. F. Naylor, ''On visible surface generation by a priori tree structures,'' *ACM Computer Graphics, SIGGRAPH '80 Proceedings*, July 1980, Vol. 14(3), pp. 124–133.
[54] B. F. Naylor, ''Constructing good partitioning trees,'' *Graphics Interface'93*, May 1993, pp. 181–191.
[55] J. B. Keller, ''Geometrical theory of diffraction,'' J. Opt. Soc. Am. **52**(2), 116–130 (1962).
[56] S. Teller, ''Computing the antiumbra cast by an area light source,'' *ACM Computer Graphics, SIGGRAPH'92 Proceedings*, July 1992, Vol. 26(2), pp. 139–148.
[57] J. Goodman and J. O'Rourke (eds.), *Handbook of Discrete and Computational Geometry* (CRC, Boca Raton, FL, 1997).
[58] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C, 2nd edition* (Cambridge U.P., New York, 1992).
[59] L. Aveneau, Y. Pousset, R. Vauzelle, and M. Méria ux, ''Development and evaluations of physical and computer optimizations for the 3d utd model,'' *AP2000 Millennium Conference on Antennas & Propagation* (Poster), April 2000.
[60] H. Lehnert and J. Blauert, ''Principles of binaural room simulation,'' Appl. Acoust. **36**, 259–291 (1992).
[61] M. E. Delany and E. N. Bazley, ''Acoustical characteristics of fibrous absorbent materials,'' Technical Report NPL AERO REPORT Ac37, National Physical Laboratory, Aerodynamics Division, March 1969.
[62] A. D. Pierce, *Acoustics. An introduction to its physical principles and applications*, 3rd ed. (American Institute of Physics, New York, 1984).
[63] S.-I. Thomasson, ''Reflection of waves from a point source by an impedance boundary,'' J. Acoust. Soc. Am. **59**, 780–785 (1976).
[64] R. G. Kouyoumjian and P. H. Pathak, ''A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface,'' Proc. IEEE **62**, 1448–1461 (1974).
[65] T. Kawai, ''Sound diffraction by a many sided barrier or pillar,'' J. Sound Vib. **79**(2), 229–242 (1981).
[66] M. A. Biot and I. Tolstoy, ''Formulation of wave propagation in infinite media by normal coordinates with an application to diffraction,'' J. Acoust. Soc. Am. **29**, 381–391 (1957).
[67] H. Medwin, E. Childs, and G. Jebsen, ''Impulse studies of double diffraction: A discrete huygens interpretation,'' J. Acoust. Soc. Am. **72**, 1005–1013 (1982).
[68] U. P. Svensson, R. I. Fred, and J. Vanderkooy, ''Analytic secondary source model of edge diffraction impulse responses,'' J. Acoust. Soc. Am. **106**, 2331–2344 (1999).

[69] R. R. Torres, U. P. Svensson, and M. Kleiner, ''Computation of edge diffraction for more accurate room acoustics auralization,'' J. Acoust. Soc. Am. **109**, 600–610 (2001).
[70] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge, MA, 1983).
[71] A. Stettner and D. P. Greenberg, ''Computer graphics visualization for acoustic simulation,'' *ACM Computer Graphics, SIGGRAPH'89 Proceedings*, July 1989, Vol. 23(3), pp. 195–206.
[72] J. Arvo and D. Kirk, ''A survey of ray tracing acceleration techniques,'' *An Introduction to Ray Tracing*, 1989.
[73] C. Sequin and E. Smyrl, ''Parameterized ray tracing,'' *ACM Computer Graphics, SIGGRAPH'89 Proceedings*, July 1989, Vol. 23(3), pp. 307–314.
[74] M. Monks, B. M. Oh, and J. Dorsey, ''Audioptimization: Goal based acoustic design,'' *IEEE Computer Graphics & Applications*, May 2000, pp. 76–91.
[75] M. Vorlander, ''International round robin on room acoustical computer simulations,'' in *Proceedings of the 15th International Congress of Acoustics*, June 1995.
[76] K. Nakagawa, T. Miyajima, and Y. Tahara, ''An improved geometrical sound field analysis in rooms using scattered sound and an audible room acoustic simulator,'' J. Appl. Acoust. **38**(2-4), 115–130 (1993).
[77] G. M. Naylor and J. H. Rindel, ''Predicting room acoustical behavior with the odeon computer model,'' in *Proceedings of the 124th ASA Meeting*, November 1992, p. 3aAA3.

Wideband RELAX and wideband CLEAN for aeroacoustic imaging

Yanwei Wang and Jian Li[a)]
Department of Electrical and Computer Engineering, University of Florida, Gainesville, Florida 32611

Petre Stoica
Department of Systems and Control, Uppsala University, P.O. Box 337, SE-75105 Uppsala, Sweden

Mark Sheplak
Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, Florida 32611

Toshikazu Nishida
Department of Electrical and Computer Engineering, P.O. Box 116200, University of Florida, Gainesville, Florida 32611

(Received 18 April 2003; revised 3 November 2003; accepted 17 November 2003)

Microphone arrays can be used for acoustic source localization and characterization in wind tunnel testing. In this paper, the wideband RELAX (WB-RELAX) and the wideband CLEAN (WB-CLEAN) algorithms are presented for aeroacoustic imaging using an acoustic array. WB-RELAX is a parametric approach that can be used efficiently for point source imaging without the sidelobe problems suffered by the delay-and-sum beamforming approaches. WB-CLEAN does not have sidelobe problems either, but it behaves more like a nonparametric approach and can be used for both point source and distributed source imaging. Moreover, neither of the algorithms suffers from the severe performance degradations encountered by the adaptive beamforming methods when the number of snapshots is small and/or the sources are highly correlated or coherent with each other. A two-step optimization procedure is used to implement the WB-RELAX and WB-CLEAN algorithms efficiently. The performance of WB-RELAX and WB-CLEAN is demonstrated by applying them to measured data obtained at the NASA Langley Quiet Flow Facility using a small aperture directional array (SADA). Somewhat surprisingly, using these approaches, not only were the parameters of the dominant source accurately determined, but a highly correlated multipath of the dominant source was also discovered.
[DOI: 10.1121/1.1639906]

PACS numbers: 43.60.Fg, 43.60.Jn, 43.60.Lq [EJS] Pages: 757–767

I. INTRODUCTION

As aircraft noise regulations become more stringent, manufacturers seek methods to reduce the noise generated by an aircraft. With the development of quieter engines, the noise generated by the airframe becomes more significant, particularly during the approaching and landing phases of a flight (Bent *et al.*, 1996). To design quieter aircraft, the physical mechanisms of noise generation must be understood. Localizing a noise source and analyzing the spatial and temporal characteristics of its radiation can provide insight into the noise generation mechanisms and aid in constructing theoretical or computational noise generation models (Arnold, 2001).

Microphone array processing methods have been developed for the purpose of obtaining accurate spatial and temporal information about acoustic sources (Humphreys *et al.*, 1998; Piet and Elias, 1997). Different aspects of array processing have been discussed in the literature, including, for example, array topology (Piet and Elias, 1997), array calibration (Mosher *et al.*, 1997), array beamforming algorithms (Dougherty and Stoker, 1998; Gramann and Mocio, 1993), data processing and visualization systems (Watts *et al.*, 1996; Mosher *et al.*, 1999), and array testing (Humphreys *et al.*, 1998; Brooks and Humphreys, 1999; Mosher, 1996; Piet *et al.*, 1999; Meadows *et al.*, 1997; Storms *et al.*, 1999).

To form high-quality acoustic images using microphone arrays, both delay-and-sum (Humphreys *et al.*, 1998) and adaptive (Dougherty and Stoker, 1998; Gramann and Mocio, 1993) beamforming methods have been considered. Besides the high sidelobes, one inherent problem of the delay-and-sum beamforming method for wideband signals is the varying mainlobe width as a function of frequency when a wideband signal is divided into many narrow-band frequency bins via the Fourier transform. One approach to mitigate this problem is to use a special shading or weighting algorithm with the delay-and-sum method (Brooks *et al.*, 1987).

The adaptive algorithms (Dougherty and Stoker, 1998; Gramann and Mocio, 1993) have much better resolution and much better interference suppression capability than the delay-and-sum approaches. However, they also suffer from the varying mainlobe width problem as the frequency changes. Moreover, they exhibit severe performance degra-

[a)]Author to whom correspondence should be addressed. Current address: Department of Electrical and Computer Engineering, P.O. Box 116130, University of Florida, Gainesville, FL 32611. Electronic mail: li@dsp.ufl.edu

dations when the number of snapshots is small and the sources are highly correlated or coherent (Van Trees, 2002).

In this paper, we present two closely related algorithms, namely wideband RELAX (WB-RELAX) and wideband CLEAN (WB-CLEAN), for aeroacoustic imaging using a near-field acoustic array. The former is a parametric approach that can be used efficiently for point source imaging without the sidelobe problems suffered by the delay-and-sum type of approaches. The latter does not have sidelobe problems either but behaves more like a nonparametric approach and it can be used for both point source and distributed source imaging. Moreover, neither algorithm suffers from the severe performance degradations encountered by the adaptive beamforming methods described above. Furthermore, WB-RELAX and WB-CLEAN eliminate the varying mainlobe width problem suffered by the delay-and-sum approaches. A two-step optimization procedure is used to implement the WB-RELAX and WB-CLEAN algorithms efficiently. The performance of WB-RELAX and WB-CLEAN is demonstrated by applying them to measured data obtained at the NASA Langley Quiet Flow Facility using a small aperture directional array (SADA) (Humphreys *et al.*, 1998; Brooks and Humphreys, 1999). Somewhat surprisingly, using these algorithms, we have not only been able to accurately determine the parameters of the dominant source, but also discovered a highly correlated multipath of the dominant source. The correlation between the dominant source and its multipath is analyzed and a conjecture on the possible cause of the multipath is provided. Three-dimensional (3-D) images of the acoustic sources are also constructed using WB-RELAX and WB-CLEAN with the NASA measured data.

The remainder of this paper is organized as follows. In Sec. II, we describe the wideband signal model and formulate the aeroacoustic imaging problem as a signal parameter estimation problem. In Secs. III and IV, we present the WB-RELAX and WB-CLEAN algorithms, respectively. Implementation issues and experimental results are discussed in Sec. V. Finally, we conclude the paper in Sec. VI.

II. PROBLEM FORMULATION

Consider an M-element microphone array receiving the wavefield generated by K near-field wideband sources in the presence of additive noise. For each microphone output, the time-domain sampled data sequence is divided into I nonoverlapping segments with each segment containing L samples. The L-point discrete Fourier transform (DFT) is applied to each segment to transform the wideband data into L narrow-band frequency bins. [Note that a sufficiently large L should be chosen so that the narrow-band assumption (Zatman, 1998) is valid.] According to the frequency-domain snapshot model (Van Trees, 2002), the $M\times 1$ noisy array output vector $\mathbf{y}_i(\omega_l)$ at the lth frequency bin of the ith segment can be written as

$$\mathbf{y}_i(\omega_l)=\mathbf{A}(\mathbf{p},\omega_l)\mathbf{s}_i(\omega_l)+\mathbf{e}_i(\omega_l),\quad l=1,...,L,\quad i=1,...,I, \tag{1}$$

where $\mathbf{s}_i(\omega_l)$ is the $K\times 1$ signal vector $\mathbf{s}_i(\omega_l)=[s_{i,1}(\omega_l)\, s_{i,2}(\omega_l)\cdots s_{i,K}(\omega_l)]^T$, $\mathbf{e}_i(\omega_l)$ is the $M\times 1$ noise vector, and the $M\times K$ steering matrix is $\mathbf{A}(\mathbf{p},\omega_l)\triangleq[\mathbf{a}(p_1,\omega_l)\,\mathbf{a}(p_2,\omega_l)\cdots\mathbf{a}(p_K,\omega_l)]$ with $\mathbf{p}\triangleq[p_1,p_2,\ldots,p_K]^T$ being the vector of the 3-D locations p_k, $k=1,...,K$, of the incident signals. The $M\times 1$ kth steering vector $\mathbf{a}(p_k,\omega_l)$ has the following form:

$$\mathbf{a}(p_k,\omega_l)=\begin{bmatrix}\frac{1}{r_{1,k}}e^{-j\omega_l r_{1,k}/c_0}\\ \vdots\\ \frac{1}{r_{m,k}}e^{-j\omega_l r_{m,k}/c_0}\\ \vdots\\ \frac{1}{r_{M,k}}e^{-j\omega_l r_{M,k}/c_0}\end{bmatrix}, \tag{2}$$

where $r_{m,k}$ is the distance between the mth microphone and the kth source, and c_0 is the (known) sound propagation velocity.

In our formulation, the noise vector $\mathbf{e}_i(\omega_l)$, $l=1,...,L$, is assumed to be a zero-mean spatially correlated but "temporally" (versus look i and sample l) white random process. The signal waveform vector $\mathbf{s}_i(\omega_l)$ is modeled as a deterministic unknown. The problem of interest herein is to estimate $\mathbf{s}_i(\omega_l)$ and $\mathbf{p}$ from the measured data $\mathbf{y}_i(\omega_l)$, $l=1,...,L$, $i=1,...,I$. For the methods studied here, we can choose $I=1$, but for adaptive beamforming, $I>M$ is required in general. Once we have estimated the parameters of the incident sources, the corresponding 3-D acoustic image can be constructed based on these results.

III. WIDEBAND RELAX

The RELAX algorithm was first proposed by Li and Stoica (1996) for mixed spectrum estimation. Later, it was extended by Li *et al.* (1997) to angle-of-arrival and waveform estimation of wideband plane waves arriving at a uniform linear array (ULA) under the spatially correlated noise assumption. In this section, we present the wideband RELAX (WB-RELAX) algorithm as a parameter estimation approach to estimate the locations and the waveforms of near-field wideband sources using an arbitrary acoustic array.

We consider the wideband signal parameter estimation problem formulated in Sec. II. WB-RELAX estimates the signal parameters $\{\mathbf{p},\mathbf{s}_i(\omega_l)\}$ in the data model given in Eq. (1) by minimizing the following nonlinear least-squares criterion:

$$G_r=\sum_{i=1}^{I}\sum_{l=1}^{L}[\mathbf{y}_i(\omega_l)-\mathbf{A}(\mathbf{p},\omega_l)\mathbf{s}_i(\omega_l)]^H\times[\mathbf{y}_i(\omega_l)-\mathbf{A}(\mathbf{p},\omega_l)\mathbf{s}_i(\omega_l)], \tag{3}$$

where the subscript "r" stands for RELAX and $(\cdot)^H$ denotes the conjugate transpose. If the noise vector $\mathbf{e}_i(\omega_l)$ were both spatially and temporally white and Gaussian, minimizing Eq. (3) would yield the maximum likelihood estimates of the source locations and the corresponding source waveforms. Even when the noise vector $\mathbf{e}_i(\omega_l)$ is spatially colored, as assumed here, it has been shown by Stoica and Nehorai

(1989) and Stoica *et al.* (1997) that the estimates of $\{\mathbf{p},\mathbf{s}_i(\omega_l)\}$ obtained by minimizing Eq. (3) are still asymptotically (for $M \gg 1$) statistically efficient under mild conditions. That is, these estimates asymptotically achieve the Cramér–Rao bound (CRB), which is the best performance bound for an unbiased estimator, in spite of the fact that $\mathbf{e}_i(\omega_l)$ is not spatially white.

To minimize G_r, consider first minimizing with respect to $\mathbf{s}_i(\omega_l)$ for a fixed $\mathbf{p}$. This yields the well-known solution

$$\hat{\mathbf{s}}_i(\omega_l) = (\mathbf{A}^H(\mathbf{p},\omega_l)\mathbf{A}(\mathbf{p},\omega_l))^{-1}\mathbf{A}^H(\mathbf{p},\omega_l)\mathbf{y}_i(\omega_l), \quad i=1,...,I, \quad l=1,...,L. \tag{4}$$

Substituting Eq. (4) into Eq. (3) gives the following minimization problem:

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} \sum_{i=1}^{I}\sum_{l=1}^{L} \|[\mathbf{I}-\mathbf{A}(\mathbf{p},\omega_l) \times (\mathbf{A}^H(\mathbf{p},\omega_l)\mathbf{A}(\mathbf{p},\omega_l))^{-1}\mathbf{A}^H(\mathbf{p},\omega_l)]\mathbf{y}_i(\omega_l)\|^2, \tag{5}$$

where $\|\cdot\|$ denotes the Euclidean norm. The minimization of Eq. (5) is a highly nonlinear $3K$-dimensional optimization problem, which we will try to avoid. The WB-RELAX algorithm does not concentrate out $\hat{\mathbf{s}}_i(\omega_l)$ but rather decouples the $3K$-dimensional optimization problem in Eq. (3) into a sequence of 3-D optimization problems in a conceptually and computationally simple way.

Before presenting our WB-RELAX algorithm, we first provide the following preparations. Assume that there are $\bar{K}$ incident signals and we have estimated the parameters $\hat{p}_j$, $\hat{s}_{i,j}(\omega_l)$ of all but the kth signal. Let

$$\mathbf{y}_{k,i}(\omega_l) \triangleq \mathbf{y}_i(\omega_l) - \sum_{j=1,j\neq k}^{\bar{K}} \mathbf{a}(\hat{p}_j,\omega_l)\hat{s}_{i,j}(\omega_l), \quad i=1,...,I, \quad l=1,...,L. \tag{6}$$

Then the least-squares cost function for the kth signal is given by

$$G_k = \sum_{i=1}^{I}\sum_{l=1}^{L} [\mathbf{y}_{k,i}(\omega_l)-\mathbf{a}(p_k,\omega_l)s_{i,k}(\omega_l)]^H \times [\mathbf{y}_{k,i}(\omega_l)-\mathbf{a}(p_k,\omega_l)s_{i,k}(\omega_l)]. \tag{7}$$

Minimizing G_k with respect to $s_{i,k}(\omega_l)$ and p_k, we have

$$\hat{s}_{i,k}(\omega_l) = \left.\frac{\mathbf{a}^H(p_k,\omega_l)\mathbf{y}_{k,i}(\omega_l)}{\mathbf{a}^H(p_k,\omega_l)\mathbf{a}(p_k,\omega_l)}\right|_{p_k=\hat{p}_k} \tag{8}$$

and

$$\hat{p}_k = \arg\min_{p_k} \sum_{i=1}^{I}\sum_{l=1}^{L} \left\| \left[\mathbf{I} - \frac{\mathbf{a}(p_k,\omega_l)\mathbf{a}^H(p_k,\omega_l)}{\mathbf{a}^H(p_k,\omega_l)\mathbf{a}(p_k,\omega_l)}\right]\mathbf{y}_{k,i}(\omega_l)\right\|^2 \tag{9}$$

$$= \arg\max_{p_k} \sum_{i=1}^{I}\sum_{l=1}^{L} \frac{|\mathbf{a}^H(p_k,\omega_l)\mathbf{y}_{k,i}(\omega_l)|^2}{\|\mathbf{a}(p_k,\omega_l)\|^2}. \tag{10}$$

We note from Eq. (10) that $\hat{p}_k$ is the location of the dominant peak of the sum over (i,l) of the normalized delay-and-sum beamformers $|\mathbf{a}^H(p_k,\omega_l)\mathbf{y}_{k,i}(\omega_l)|^2/\|\mathbf{a}(p_k,\omega_l)\|^2$. Equation (10) is also recognized as a matched filtering of the received wavefield, $\mathbf{a}(p_k,\omega_l)s_{i,k}(\omega_l)$, under the assumption that $\mathbf{y}_{k,i}(\omega_l)$ contains only the kth signal and additive noise. Now we are ready to present the WB-RELAX algorithm.

Step 1: Assume $\bar{K}=1$. Obtain $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=1}$ by using Eqs. (10) and (8) with $\mathbf{y}_{1,i}(\omega_l)=\mathbf{y}_i(\omega_l)$.

Step 2: Assume $\bar{K}=2$. Compute $\mathbf{y}_{2,i}(\omega_l)$ with Eq. (6) by using $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=1}$ obtained in **Step 1**. Obtain $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=2}$ from $\mathbf{y}_{2,i}(\omega_l)$. Next, compute $\mathbf{y}_{1,i}(\omega_l)$ by using $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=2}$ and then redetermine $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=1}$ from $\mathbf{y}_{1,i}(\omega_l)$. Iterate the previous two substeps until practical convergence is achieved (to be discussed later on).

Step 3: Assume $\bar{K}=3$. Compute $\mathbf{y}_{3,i}(\omega_l)$ by using $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=1,2}$ obtained in **Step 2**. Obtain $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=3}$ from $\mathbf{y}_{3,i}(\omega_l)$. Next, compute $\mathbf{y}_{1,i}(\omega_l)$ by using $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=2,3}$, and redetermine $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=1}$ from $\mathbf{y}_{1,i}(\omega_l)$. Then, compute $\mathbf{y}_{2,i}(\omega_l)$ by using $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=1,3}$ and redetermine $\{\hat{p}_k,\hat{s}_{i,k}(\omega_l)\}_{k=2}$ from $\mathbf{y}_{2,i}(\omega_l)$. Iterate these substeps until practical convergence.

Remaining Steps: Continue similarly until $\bar{K}$ is equal to K, which is assumed known or empirically determined.

When the sources are not spaced very closely, i.e., not less than half of the delay-and-sum beamformer's resolution limit, WB-RELAX usually converges in a few steps (Li and Stoica, 1996). Otherwise, WB-RELAX may converge slowly. The "practical convergence" of the WB-RELAX algorithm may be determined by checking the relative change, say lower than $\epsilon = 10^{-3}$, of the cost function G_r in Eq. (3) between two consecutive iterations.

The kth wideband signal waveform estimate can be calculated from the L-point inverse discrete Fourier transform (IDFT) of $\{\hat{s}_{i,k}(\omega_l)\}_{l=1}^{L}$. Also we estimate the signal power using the following equation:

$$\hat{P}_k = \frac{1}{I}\sum_{i=1}^{I}\sum_{l=1}^{L} |\hat{s}_{i,k}(\omega_l)|^2. \tag{11}$$

Based on the descriptions above, the WB-RELAX algorithm can be summarized as follows:

WB-RELAX Algorithm
Let $\bar{K}=1$, $\mathbf{y}_{1,i}(\omega_l)=\mathbf{y}_i(\omega_l)$, for $l=1,...,L$, and $i=1,...,I$
$\hat{p}_1 = \arg\max_{p_1} \Sigma_{i=1}^{I}\Sigma_{l=1}^{L}|\mathbf{a}^H(p_1,\omega_l)\mathbf{y}_{1,i}(\omega_l)|^2/\|\mathbf{a}(p_1,\omega_l)\|^2$
$\hat{s}_{i,1}(\omega_l) = \mathbf{a}^H(\hat{p}_1,\omega_l)\mathbf{y}_{1,i}(\omega_l)/\|\mathbf{a}(\hat{p}_1,\omega_l)\|^2$ for $l=1,...,L$, and $i=1,..,I$
for $\bar{K}=2$ to K **do**
  repeat
    for $k=\bar{K},1,2,3,...,\bar{K}-1$ **do**
$\mathbf{y}_{k,i}(\omega_l) = \mathbf{y}_i(\omega_l) - \Sigma_{j=1,j\neq k}^{\bar{K}}\mathbf{a}(\hat{p}_j,\omega_l)\hat{s}_{i,j}(\omega_l)$ for $l=1,...,L$, and $i=1,...,I$

$\hat{p}_k = \arg\max_{p_k} \Sigma_{i=1}^{I}\Sigma_{l=1}^{L}|\mathbf{a}^H(p_k,\omega_l)\mathbf{y}_{k,i}(\omega_l)|^2/\|\mathbf{a}(p_k,\omega_l)\|^2$
$\hat{s}_{i,k}(\omega_l) = \mathbf{a}^H(\hat{p}_k,\omega_l)\mathbf{y}_{k,i}(\omega_l)/\|\mathbf{a}(\hat{p}_k,\omega_l)\|^2$ for $l=1,...,L$, and $i=1,...,I$
    end for
  until practical convergence achieved
end for

IV. WIDEBAND CLEAN

The CLEAN algorithm was first proposed in radio astronomy (Högbom, 1974) for gapped data (or irregularly sampled data) analysis. In the original CLEAN algorithm (Högbom, 1974), it was assumed that the target space can be represented by a small number of point sources in an otherwise empty field of view. Later, CLEAN was applied to the case of distributed sources with a small loop gain parameter. From the aperture synthesis theory, we know that the synthesis map obtained via the delay-and-sum beamformer is proportional to the convolution of the true source intensity distribution with the synthesized beam. Because of the finite number of samples, the data gaps, and the irregularity of the data sampling process, the synthesized beam (the so-called ''dirty beam'' (Högbom, 1974)) will comprise an extended pattern of undesirable sidelobes. As a consequence, the synthesis map convolving the source intensity distribution with the synthesized beam will be disturbed by the sidelobes from the synthesis beam, and the result is called the ''dirty map'' (Högbom, 1974). The CLEAN algorithm attempts to deconvolve the synthesized map into the true source intensity distribution and the ''dirty beam'' using an iterative method.

In this section, we extend the CLEAN algorithm to the wideband signal and multisnapshot case. At the *n*th iteration, to estimate the most significant signal impinging on the microphone array, we consider the following least-squares criterion:

$$G_c=\sum_{i=1}^{I}\sum_{l=1}^{L}[\mathbf{y}_{i,n}(\omega_l)-\mathbf{a}(p_n,\omega_l)s_{i,n}(\omega_l)]^H \times[\mathbf{y}_{i,n}(\omega_l)-\mathbf{a}(p_n,\omega_l)s_{i,n}(\omega_l)], \tag{12}$$

where the subscript ''*c*'' stands for CLEAN and

$$\mathbf{y}_{i,n}(\omega_l)=\mathbf{y}_i(\omega_l)-\sum_{\tilde{n}=1}^{n-1}\rho\cdot\mathbf{a}(\hat{p}_{\tilde{n}},\omega_l)\hat{s}_{i,\tilde{n}}(\omega_l), \tag{13}$$

with $\hat{p}_{\tilde{n}}$ and $\hat{s}_{i,\tilde{n}}(\omega_l)$, $\tilde{n}=1,2,...,n-1$, being obtained from the previous iterations and ρ being the loop gain usually between 0.1 and 0.25. By minimizing the above least-squares cost function with respect to $s_{i,n}(\omega_l)$ and p_n, we have

$$\hat{s}_{i,n}(\omega_l)=\left.\frac{\mathbf{a}^H(p_n,\omega_l)\mathbf{y}_{i,n}(\omega_l)}{\mathbf{a}^H(p_n,\omega_l)\mathbf{a}(p_n,\omega_l)}\right|_{p_n=\hat{p}_n} \tag{14}$$

and

$$\hat{p}_n=\arg\min_{p_n}\sum_{i=1}^{I}\sum_{l=1}^{L}\left\|\left[\mathbf{I}-\frac{\mathbf{a}(p_n,\omega_l)\mathbf{a}^H(p_n,\omega_l)}{\mathbf{a}^H(p_n,\omega_l)\mathbf{a}(p_n,\omega_l)}\right]\mathbf{y}_{i,n}(\omega_l)\right\|^2 \tag{15}$$

$$=\arg\max_{p_n}\sum_{i=1}^{I}\sum_{l=1}^{L}\frac{|\mathbf{a}^H(p_n,\omega_l)\mathbf{y}_{i,n}(\omega_l)|^2}{\|\mathbf{a}(p_n,\omega_l)\|^2}. \tag{16}$$

Note from Eqs. (16) and (14) that the above estimates of the source location and the signal waveform are similar to Eqs. (10) and (8) of WB-RELAX, respectively.

The WB-CLEAN algorithm can now be described as follows.

Step 1: Choose ρ, initialize with $n=1$ and let $\mathbf{y}_{i,1}(\omega_l)=\mathbf{y}_i(\omega_l)$.

Step 2: Estimate p_n and $s_{i,n}(\omega_l)$ by using Eqs. (16) and (14), for $l=1,2,...,L$, $i=1,...,I$.

Step 3: Calculate the array received wavefield generated by the source estimated in the last step. Subtract out only a fraction ρ (e.g., $\rho=0.1$) of the estimated wavefield from $\mathbf{y}_{i,n}(\omega_l)$ to get $\mathbf{y}_{i,n+1}(\omega_l)$.

Step 4: Add the current estimated signal location and waveform (scaled by ρ) to a list of estimated signals.

Step 5: Increase the iteration index n by 1.

Remaining Steps: Repeat Steps 2–5 until the process works down to the noise level.

We terminate our iterations by calculating the power distribution $D_P^n(p)$, where

$$D_P^n(p)=\sum_{i=1}^{I}\sum_{l=1}^{L}\frac{|\mathbf{a}^H(p,\omega_l)\mathbf{y}_{i,n}(\omega_l)|^2}{\|\mathbf{a}(p,\omega_l)\|^2}, \tag{17}$$

and its entropy $H(D_P^n(p))$, where

$$H(X)=-\sum_{x\in\chi}f(x)\log f(x) \tag{18}$$

and $f(x)$ is the probability mass function (pmf) of X.

In our examples in Sec. V, we terminate the iterations, based on our empirical experience, when the entropy is around 4.35.

We also note that the CLEAN algorithm subtracts out a small fraction ρ of the ''dirty beam'' from the ''dirty map'' at each iteration. Another form of the CLEAN algorithm, proposed by Dougherty and Stoker (1998), is based on covariance fitting, but it cannot deal with correlated signals or small snapshot numbers.

Based on the descriptions above, the WB-CLEAN algorithm can be summarized as follows:

WB-CLEAN Algorithm

Select the loop gain ρ

Let $n=1$, $\mathbf{y}_{i,n}(\omega_l)=\mathbf{y}_i(\omega_l)$, for $l=1,...,L$, and $i=1,...,I$

repeat

$\hat{p}_n=\arg\max_{p_n}\Sigma_{i=1}^{I}\Sigma_{l=1}^{L}|\mathbf{a}^H(p_n,\omega_l)\mathbf{y}_{i,n}(\omega_l)|^2/\|\mathbf{a}(p_n,\omega_l)\|^2$

$\hat{s}_{i,n}(\omega_l)=\mathbf{a}^H(\hat{p}_n,\omega_l)\mathbf{y}_{i,n}(\omega_l)/\|\mathbf{a}(\hat{p}_n,\omega_l)\|^2$ for $l=1,...,L$, and $i=1,...,I$

$\mathbf{y}_{i,n+1}(\omega_l)=\mathbf{y}_{i,n}(\omega_l)-\rho\cdot\mathbf{a}(\hat{p}_n,\omega_l)\hat{s}_{i,n}(\omega_l)$ for $l=1,...,L$, and $i=1,...,I$

$n=n+1$

until no significant signal is left in $\mathbf{y}_{i,n}$

$N=n-1$

Finally, reconstruct the source distribution using $\hat{p}_n$ and $\rho\hat{s}_{i,n}(\omega_l)$ for $n=1,...,N$

Note that the WB-CLEAN results $\{\hat{p}_n,\rho\hat{s}_{i,n}(\omega_l)\}_{n=1}^{N}$ can usually be divided into K groups of sources $\{g_k\}_{k=1}^{K}$ where each group stands for a distributed source. The signal power of the *k*th distributed source can be calculated using the following equation:

$$\hat{P}_k=\frac{1}{I}\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{n\in g_k}|\rho\hat{s}_{i,n}(\omega_l)|^2. \tag{19}$$

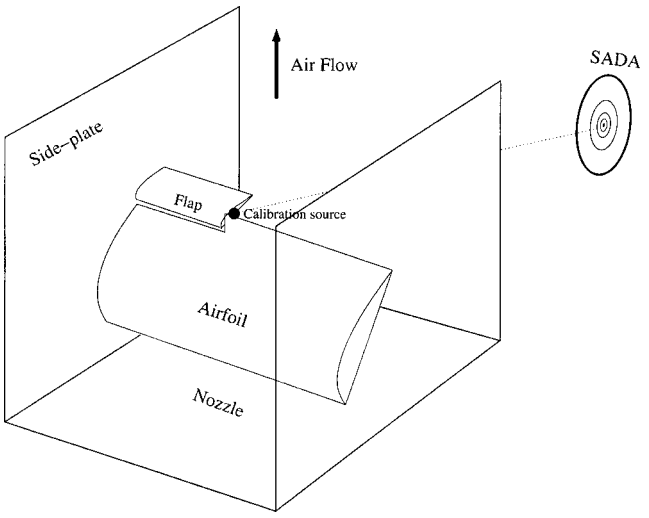


FIG. 1. Test setup for the NASA data (Humphreys *et al.*, 1998; Brooks and Humphreys, 1999).

V. EXPERIMENTAL RESULTS

To assess the algorithm performance, we have applied WB-RELAX and WB-CLEAN to the data collected in the NASA Langley Quiet Flow Facility (QFF) using a small aperture directional array (SADA) (Humphreys *et al.*, 1998; Brooks and Humphreys, 1999). The test setup is illustrated in Fig. 1. SADA is a 2-D array which has 33 microphones with a maximum diagonal aperture size of 7.76 in. The array pattern incorporates four irregular circles of eight microphones each and one microphone is placed at the array center (Brooks and Humphreys, 1999). Each circle is twice the diameter of the circle it encloses. QFF is a quiet open-jet facility designed for anechoic acoustic testing. A two-by-three feet rectangular open-jet nozzle was placed at the bottom of the QFF. A three feet span model was mounted in the QFF. The noise source configurations studied are flap edges (Brooks and Humphreys, 1999), with a calibration source placed next to the flap. The SADA was mounted on a pivotal boom designed to allow it to face the center of the flap edge of the main element airfoil (also the position of the calibration source) while maintaining a constant distance (60 in.) away from it. The test was conducted with one point source (the calibration source) but without air flow. For detailed information on the testing environment and procedure, we refer the reader to (Humphreys *et al.*, 1998; Brooks and Humphreys, 1999). The current study is limited to the no-flow case, consisting of well-defined, compact sources radiating in a homogeneous medium. An extention to trailing edge noise source studies in an open jet wind tunnel would require modification of the steering vectors to account for shear layer diffraction effects (Humphreys *et al.*, 1998).

Before applying the data, we first discuss some implementation issues of the algorithms.

A. Implementation of WB-RELAX

The most challenging part of implementing WB-RELAX (as well as WB-CLEAN, see below) is the estimation of the source location p_k, $k=1,...,K$, in Eq. (10). This maximization problem is nonconvex, and an exhaustive search method has to be used to achieve the global maximum. However, a full 3-D search is computationally intensive. In this paper, we propose a two-step search procedure which can reduce the computational complexity significantly.

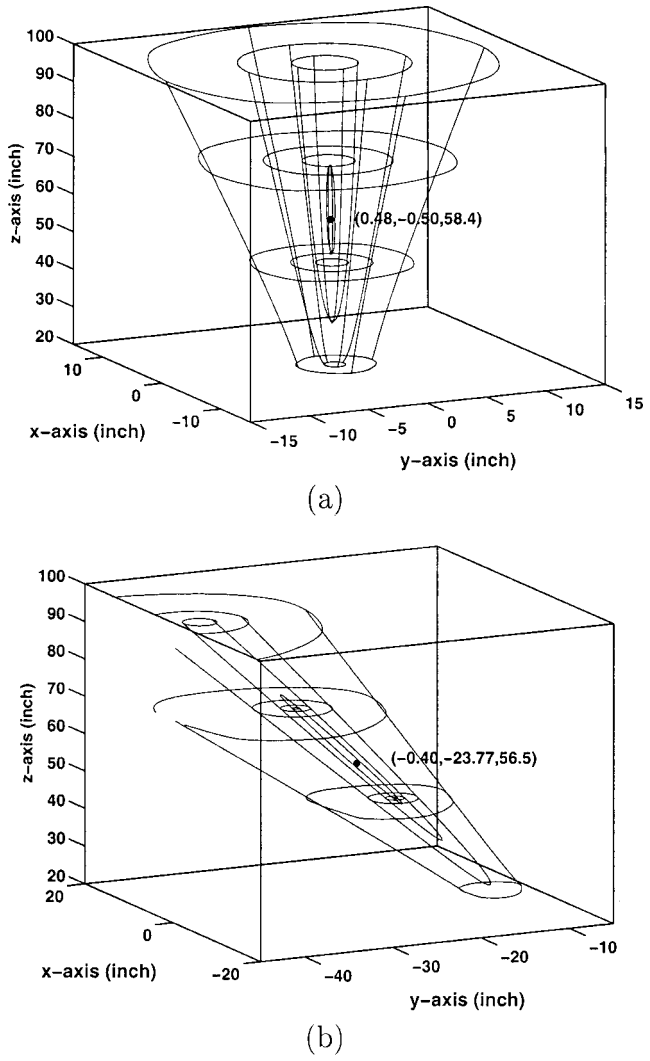


FIG. 2. Location estimation via WB-RELAX: (a) 3-D localization of the first point source; (b) 3-D localization of the second point source.

Step 1 Coarse Search: Evaluate the objective function in Eq. (10) at each point in a predefined 3-D coarse grid (i.e., 2-in. spacing for all three dimensions) within the space of interest for the NASA data. Find the position corresponding to the maximum value (coarse search result).

Step 2 Fine Search: Using the coarse search result as an initial point, perform a fine search in the vicinity of this initial point to accurately locate the location of the global maximum of Eq. (10). This can be implemented by using a derivative-free multidimensional uphill search method such as the Nelder-Mead algorithm (Nelder and Mead, 1965; Rosenbrock, 1960; Powell, 1964).

In this two-step search procedure, **Step 1** can give us the approximate position of the global maximum under the mild condition that the coarse grid is fine enough. **Step 2** can be implemented by using the ''fminsearch'' command in the Matlab optimization toolbox.

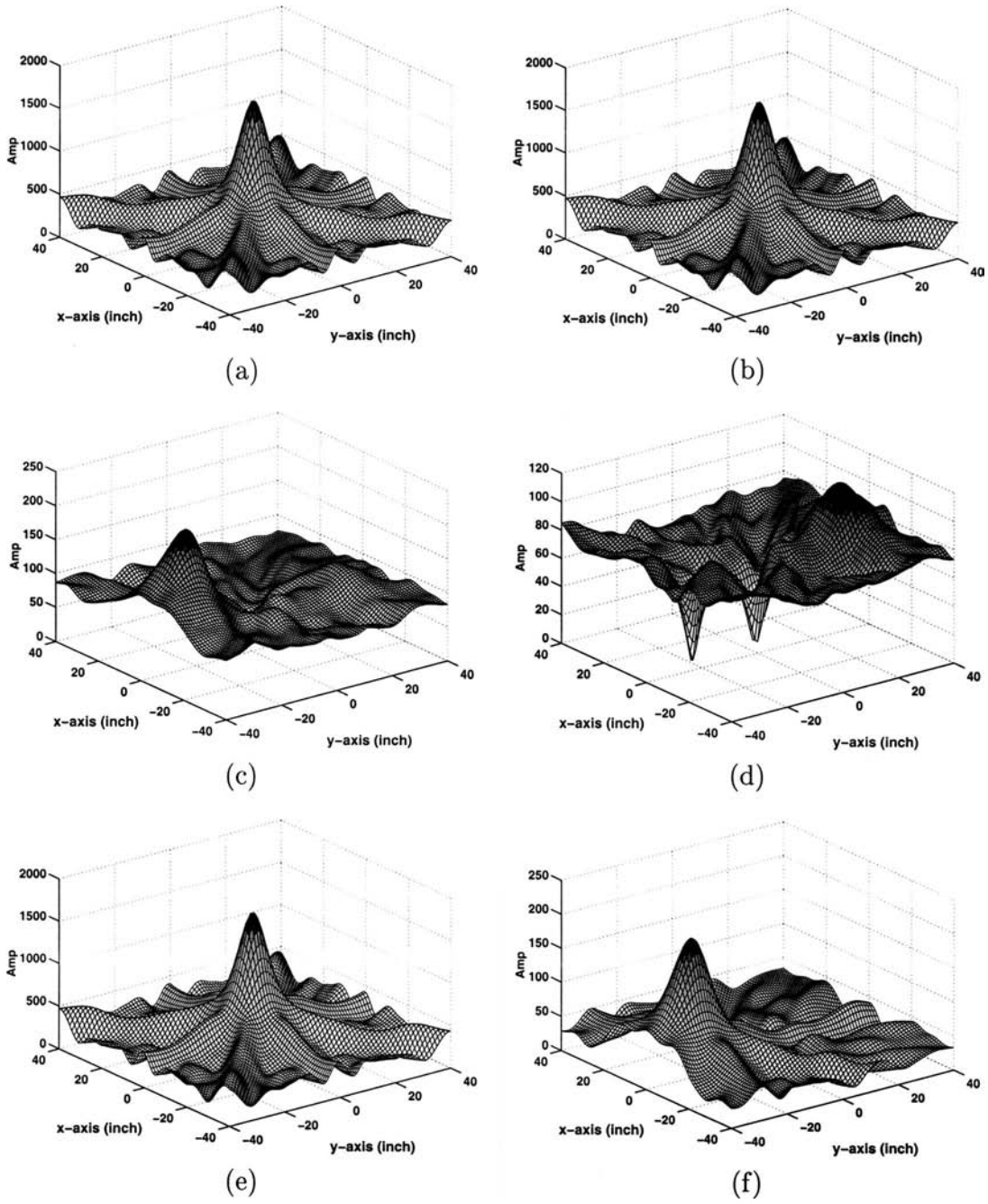


FIG. 3. Detailed results of WB-RELAX: (a) Magnitude distribution at $z=58.4$ in.. (b) Magnitude distribution at $x=58.4$ in. after subtracting out the second source. (c) Magnitude distribution at $z=56.5$ in. after subtracting out the first source. (d) Magnitude distribution of the residue data at $x=56.5$ in. after subtracting out the first and the second sources. (e) Magnitude distribution at $z=58.4$ in. for the first source only. (f) Magnitude distribution at $z=56.5$ in. for the second source only.

B. Implementation of WB-CLEAN

The implementation issues of WB-CLEAN are similar to those of WB-RELAX. For each iteration, we use the same two-step procedure to search for the location of the strongest source and estimate the corresponding signal waveform. For WB-CLEAN, the loop gain ρ is preselected. The number of iterations N may be chosen experimentally based on certain criteria such as when the estimated peak height of the current power distribution in Eq. (17) is down to the noise level. To minimize the number of iterations required by WB-CLEAN, ρ should be unity; then WB-CLEAN becomes a noniterative version of WB-RELAX. However, for distributed sources, we choose ρ to be within the interval [0.1,0.25]. Note that a small ρ (<0.1) makes the CLEAN algorithm converge very slowly and a large ρ (>0.25) makes CLEAN not suitable for distributed sources. In our experiments, we chose $\rho=0.1$.

C. Application of WB-RELAX

The test data we used herein is the first data block from test no. 510 collected in the NASA QFF (Humphreys *et al.*, 1998; Brooks and Humphreys, 1999). The sampling rate for each channel is 142.857 kHz, and there are 8192 time do-

main samples for each microphone. After applying an 8192-point DFT to the measured data, we obtain 8192 frequency bins with a resolution of 17.45 Hz. Because the signal is real-valued, only half of the frequency bins need to be considered due to the conjugate symmetry property of the Fourier transform of a real-valued signal. We utilize all of the frequency bins between 10 and 40 kHz, for a total of 1722 frequency bins. Here we assume that we only have one snapshot for each frequency bin ($I=1$). (The adaptive beamforming methods are not suitable for this case.)

We use the Cartesian coordinate system (x,y,z) in our simulations. The microphone array is located on the xy-plane, and the array center is located at (0,0,0). The space of interest is the volume between -40 and 40 in. for the x- and y-axes, and between 20 and 100 in. for the z-axis. Although only one point source was actually used in the data collection process, we chose $K=2$ sources in our WB-RELAX and discovered a second strongest source besides the dominant calibration source.

The estimated locations of the first and second sources using WB-RELAX are illustrated in Figs. 2(a) and (b), respectively. These two plots also show the 3-D contours of the search function in Eq. (10) evaluated over the space of interest for the two hypothesized sources after WB-RELAX has converged. The estimated source locations are shown by the dots in the middle of the figures, and the location coordinates are also marked correspondingly. The 3-D contour plots clearly show that SADA has a great potential for 2-D location (direction) estimation. At the same time, these two figures indicate that SADA can also be used to estimate the distances of the sources from the array center under the current testing condition, even though a bit less accurately than the estimates of the 2-D location.

Figure 3 shows more detailed results for WB-RELAX. Let

$$D_M(p)=\left[\sum_{i=1}^{I}\sum_{l=1}^{L}\frac{|\mathbf{a}^H(p,\omega_l)\tilde{\mathbf{y}}_i(\omega_l)|^2}{\|\mathbf{a}(p,\omega_l)\|^2}\right]^{1/2} \tag{20}$$

denote the magnitude distribution of the data $\tilde{\mathbf{y}}_i(\omega_l)$, which could be the original data or its processed variations, as a function of the location p. In Fig. 3(a), we show the 2-D slice of the magnitude distribution of the original data parallel to the xy-plane at $z=58.4$ in. (the estimated z-axis location of the first signal source). The dominant calibration source is shown clearly in the figure. In Fig. 3(b), the magnitude distribution at $z=58.4$ in. after coherently subtracting out the second (weaker) source from the original data is shown. In Fig. 3(c), the magnitude distribution after subtracting out the first (stronger) source is shown. Now the 2-D slice is at $z=56.5$ in., which corresponds to the estimated z-coordinate of the second source location. As we can see, after subtracting out the dominant peak in the middle of Fig. 3(a), the second source is clearly shown in Fig. 3(c). In Fig. 3(d), the magnitude distribution of the residue after subtracting out both the first and second sources, at $z=56.5$ in., is shown. It appears that there is a third weak signal left in the residue,

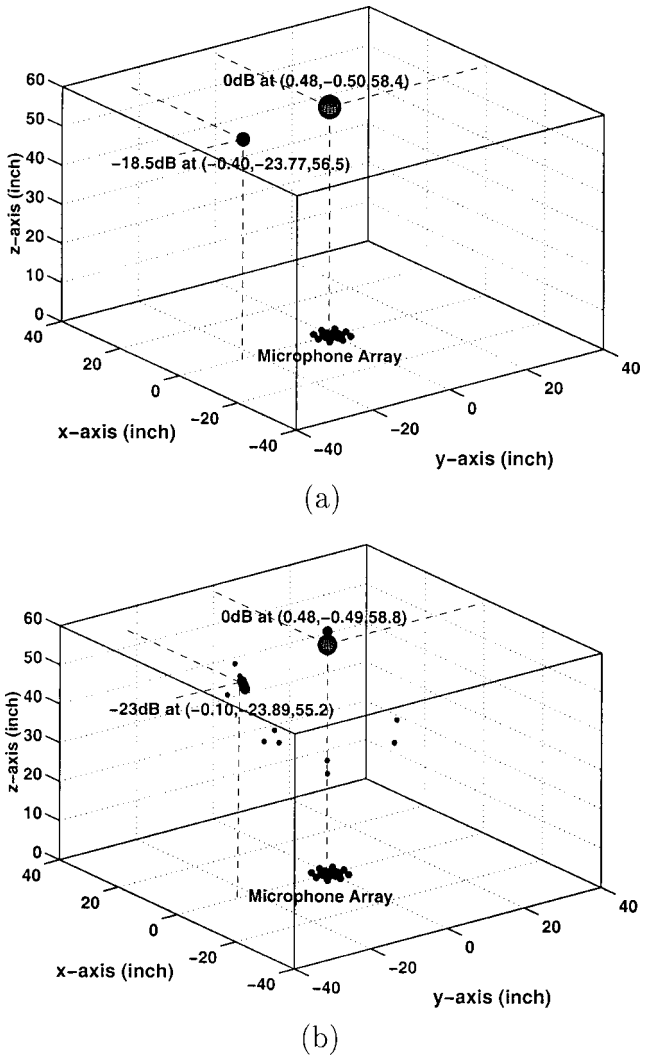


FIG. 4. Estimated locations and powers of the sources: (a) WB-RELAX; (b) WB-CLEAN.

but we did not increase K to 3 since the third signal is quite close to the noise level. The magnitude distributions of the first source only and the second source only, at their estimated z-axis locations, are shown in Figs. 3(e) and (f), respectively, which are free of noise. Note the change of scale in Fig. 3(f). Comparing Fig. 3(e) with the magnitude distribution of the original data shown in Fig. 3(a), it is clear that without WB-RELAX, we could not identify the second source, which is totally buried in the sidelobes of the calibration source. The acoustic image utilizing the results of WB-RELAX is constructed in Fig. 4(a), with the microphone array shown as well. The location coordinates and the relative powers, which are calculated by using Eq. (11), are also marked correspondingly. The estimated signal-to-noise ratio (SNR) is 14 dB.

The above results show that, although SADA is originally designed for 2-D location estimation only, we can estimate the 3-D location of the calibration source accurately using WB-RELAX and obtain the corresponding source power.

To analyze the relationship between the two estimated point sources, we need to evaluate their correlation coeffi-

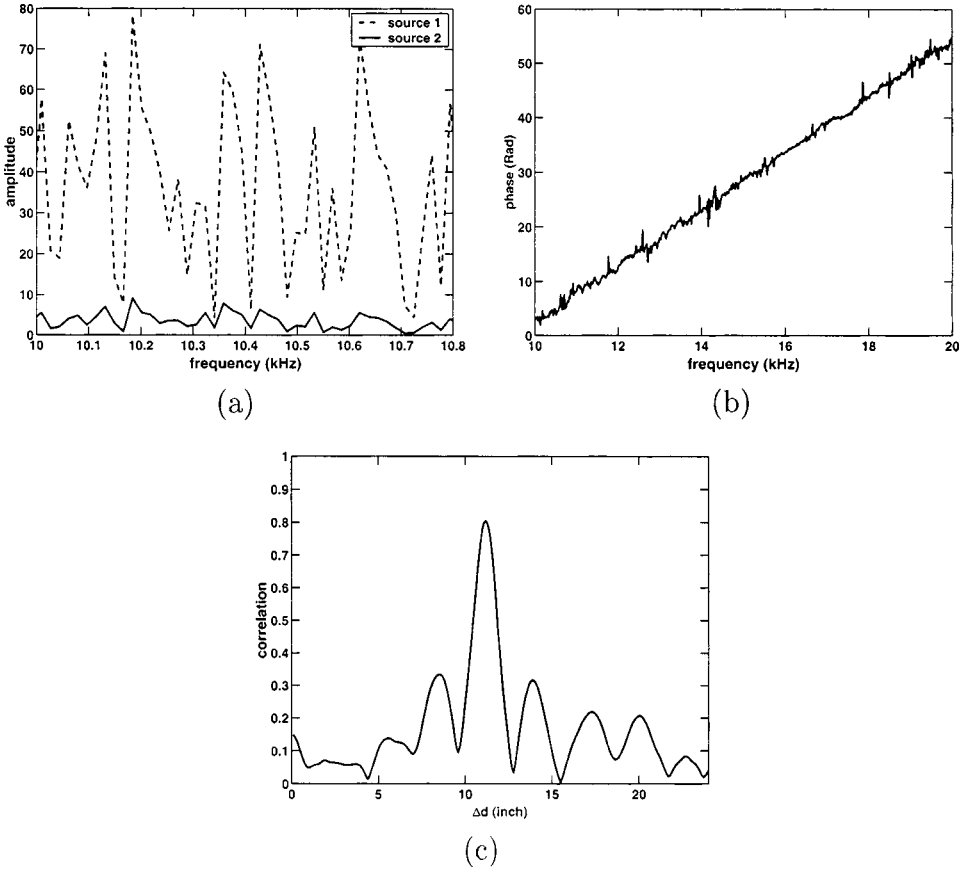


FIG. 5. Correlation between the two estimated sources via WB-RELAX: (a) Amplitude versus frequency curves for the two sources. (b) Phase difference between the two sources. (c) Magnitude of the correlation coefficient versus the propagation distance difference (delay×velocity) between the two sources.

cient. Here, we analyze the correlation for both amplitude and phase. The amplitudes of the two estimated signals over different frequency bins are plotted in Fig. 5(a). To see the relationship clearly, only a small portion of the frequency bins are included in the plot, from 10 to 10.8 kHz. The relationship is similar for other frequency bins. It can be observed that the amplitudes of the two signal waveforms are strongly correlated. The phase difference between the two estimated sources for the frequency bins from 10 to 20 kHz is plotted in Fig. 5(b), which shows clearly that the phase difference changes linearly with the frequency. Combining the results from Figs. 5(a) and (b), we conclude that the second source is a multipath of the first source [i.e., $s_2(t) = \alpha s_1(t-\tau)$, for some constants α and τ]. We calculate the correlation coefficient of the two sources over different frequency bins as a function of the propagation distance difference Δd (time delay $\tau \times$ propagation velocity c_0) between the two sources:

TABLE I. The location estimates obtained at various WB-CLEAN iterations (ρ=0.1).

Iteration	1	2	3	4	5	6	7	8	9	10	11
x-coordinate	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48
y-coordinate	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49
z-coordinate	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6
Source	1	1	1	1	1	1	1	1	1	1	1
Iteration	12	13	14	15	16	17	18	19	20	21	22
x-coordinate	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48
y-coordinate	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49	−0.49
z-coordinate	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6	58.6
Source	1	1	1	1	1	1	1	1	1	1	1
Iteration	23	24	25	26	27	28	29	30	31	32	33
x-coordinate	−0.41	0.50	−0.45	0.52	−0.49	0.59	−0.54	6.2	−0.64	−0.05	−0.87
y-coordinate	−22.7	−0.47	−22.9	−0.45	−23.2	−0.39	−23.6	4.4	−24.2	−0.89	−25.7
z-coordinate	53.4	58.9	53.9	59.4	54.6	62.0	55.6	311	57.1	29.2	60.8
Source	2	1	2	1	2	1	2		2	1	2
Iteration	34	35	36	37	38	39	40	41	42	43	44
x-coordinate	−4.06	2e6	−4.57	−4.07	−0.15	2.72	−5.13	−2e7	884	−2.7	1.3
y-coordinate	13.6	2e6	−17.0	13.7	−0.90	−24.9	−18.6	3e7	2e6	15.3	−16.3
z-coordinate	31.3	1e8	39.4	31.4	25.8	51.3	43.3	8e7	2e8	36.4	37.6
Source	3		2	3	1	2	2			3	2

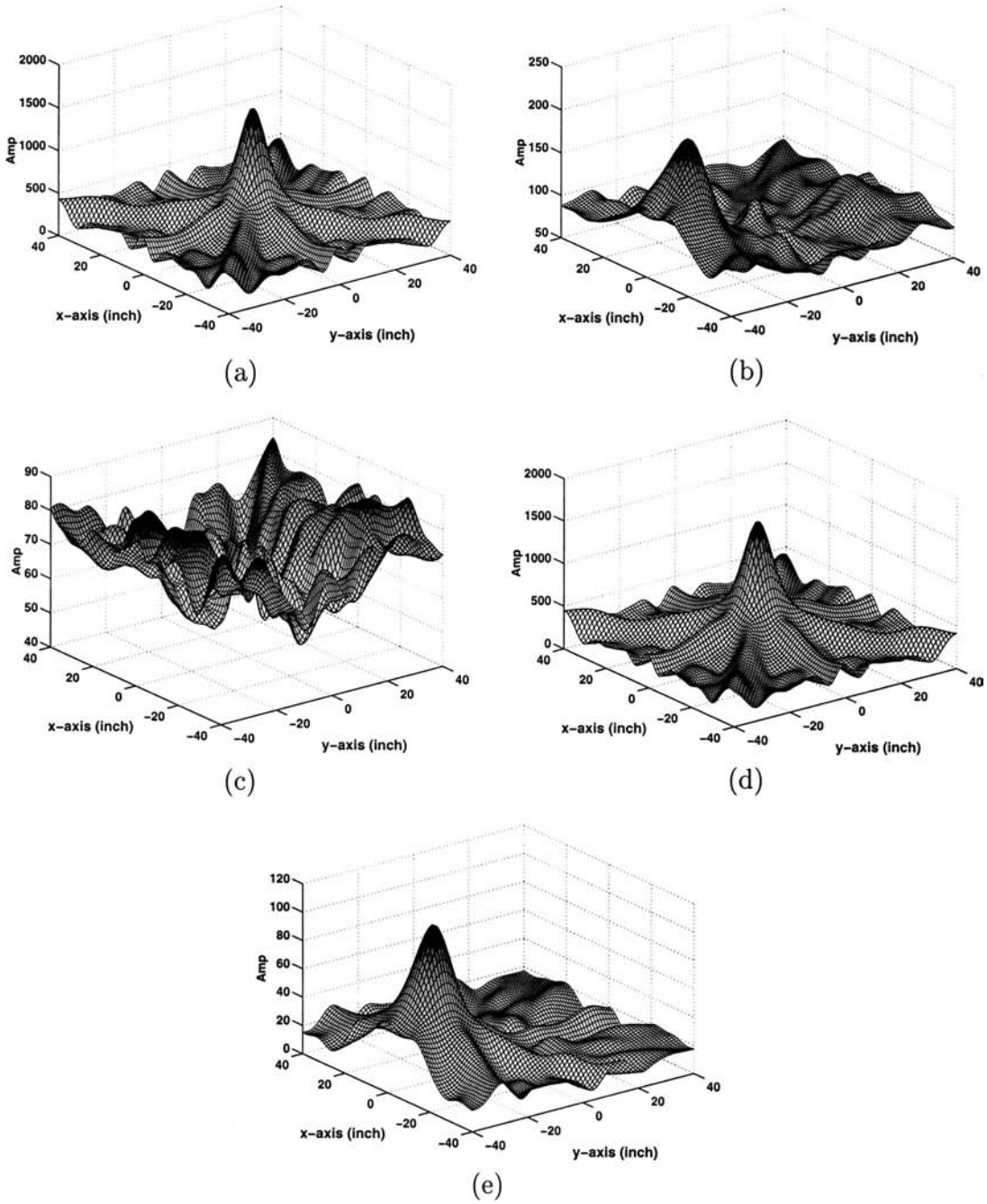


FIG. 6. Detailed results of WB-CLEAN: (a) Magnitude distribution at $z=58.8$ in. after subtracting out the second source. (b) Magnitude distribution at $z=55.2$ in. after subtracting out the first source. (c) ''Residue'' map at the first source at $z=58.8$ in.. (d) ''Clean'' map of the second source at $z=55.2$ in. (e) ''Clean'' map of $z=55.2$ in.

$$\rho_{s_1 s_2}(\Delta d) = \frac{\Sigma_{l=1}^{L} \hat{s}_{1,1}(\omega_l)\tilde{\hat{s}}^{*}_{1,2}(\omega_l)}{(\Sigma_{l=1}^{L} \hat{s}_{1,1}(\omega_l)\hat{s}^{*}_{1,1}(\omega_l))^{1/2} \cdot (\Sigma_{l=1}^{L} \tilde{\hat{s}}_{1,2}(\omega_l)\tilde{\hat{s}}^{*}_{1,2}(\omega_l))^{1/2}}, \tag{21}$$

where $(\cdot)^*$ denotes the complex conjugate and $\tilde{\hat{s}}_{1,2}(\omega_l) = \hat{s}_{1,2}(\omega_l)e^{j\Delta d/c_0\omega_l}$, $l=1,...,L$. The magnitude of $\rho_{s_1 s_2}$ is plotted in Fig. 5(c). It shows that the propagation distance of the second source is about 11 in. longer than that of the first source, and the peak correlation coefficient is over 80%. It appears that the second source is due to the reflection of the calibration source from a side-plate of QFF, which is consistent with the physical measurements.

D. Application of WB-CLEAN

We apply WB-CLEAN to the same NASA data set with $\rho=0.1$. After 44 iterations (i.e., $N=44$), there is no significant signal energy left in the residue (the estimated signal magnitude is lower than the noise level). The location estimates obtained at each iteration are given in Table I. During the first 22 iterations, the algorithm ''cleans'' the first dominant source. Starting from iteration 23, it starts to ''clean'' the second source. Note that we separate the estimates of the

TABLE II. Merits and limitations of various array processing methods.

Methods	Delay-and-sum	RELAX	CLEAN	Adaptive beamforming
Robustness	Yes	Yes	Yes	Maybe
Small snapshot no.	Yes	Yes	Yes	Maybe
Resolution	Low	High	Low	Very high
Frequency-independent resolution	No	Yes	Yes	No
Sidelobe	High	None	None	Low
Coherent sources	Yes	Yes	Yes	No
Distributed sources	Yes	No	Yes	No
Computational complexity	Low	High for large no. of sources	High for large N	High

44 iterations of WB-CLEAN into three groups: the first group corresponds to the dominant source, the second group corresponds to the second source (which is about 24 in. away from the first one along the y-axis) and the third group corresponds to noiselike sources which belong to neither the first nor the second group.

The detailed results of WB-CLEAN are shown in Fig. 6. Similar to those in Fig. 3, the surface plots of Figs. 6(a)–(e) are the 2-D magnitude distribution of different groups of sources evaluated at different z values. The magnitude distribution for each group of sources is evaluated based on the wavefield generated by all the sources within that group. Figures 6(a) and (d) correspond to $z=58.8$ in., Figs. 6(b), (c), and (e) correspond to $z=55.2$ in.

The magnitude distribution after subtracting out the second and the first sources from the original data are shown in Figs. 6(a) and (b), respectively. These results are similar to those of WB-RELAX in Figs. 3(b) and (c), respectively. The 2-D magnitude distribution of the residue is shown in Fig. 6(c). Clearly, there is no notable signal left in the residue. Comparing with Fig. 3(d), we see that WB-RELAX and WB-CLEAN give similar noise levels. The ''clean'' maps of the first and second sources are shown in Figs. 6(d) and (e), respectively. Comparing with Figs. 3(e) and (f), the peak height for the first source is almost identical with that obtained via WB-RELAX. For the second source, however, the peak height of WB-CLEAN is lower than that of WB-RELAX. This is due to the different signal estimation strategies used by WB-RELAX and WB-CLEAN. WB-RELAX estimates the signal waveform regardless of the noise level, but WB-CLEAN stops when the estimated signal power is down to the noise level. For each source, the magnitude estimated using WB-RELAX is due to both the signal and noise, while WB-CLEAN uses a cluster of small peaks above the noise level to represent the source. This is also the reason why the magnitude distribution of the residue data estimated by WB-RELAX in Fig. 3(d) has two nulls while the one estimated by WB-CLEAN in Fig. 6(c) does not. So the peak height of WB-CLEAN is always lower than that of WB-RELAX. This is more obvious for weak signals. In Fig. 4(b); the image of the acoustic source distribution is constructed using the results of WB-CLEAN. The total powers of the first and the second sources are calculated by using Eq. (19). The estimated SNR is 13.4 dB. Instead of getting separate point sources as when using WB-RELAX, WB-CLEAN gives us a map of the sources, which is a desired property if we are dealing with distributed sources. Comparing Figs. 4(a) and (b), we observe that WB-RELAX gives us a ''super clean'' image.

For comparison purposes, the merits and limitations of WB-RELAX, WB-CLEAN, as well as other existing array processing methods are listed in Table II. For the NASA data, the computational complexities of WB-RELAX and WB-CLEAN were approximately 5 and 44 times larger than that of the delay-and-sum approach.

VI. CONCLUSIONS

In this paper, two related algorithms, WB-RELAX and WB-CLEAN, were presented for the near-field acoustic imaging problem. The former is a parametric approach that can be used efficiently for point source imaging without the sidelobe problems suffered by the delay-and-sum beamforming approaches. The latter algorithm, which does not have sidelobe problems either, behaves more like a nonparametric approach and can be used for both point and distributed source imaging. Moreover, neither algorithm suffers from the severe performance degradations encountered by the adaptive beamforming methods when the number of snapshots is small and/or the sources are highly correlated or coherent with each other. The performance of WB-RELAX and WB-CLEAN was demonstrated by applying them to measured data obtained at the NASA Langley Quiet Flow Facility (QFF) using a small aperture directional array. Somewhat surprisingly, using these algorithms, we have not only accurately determined the parameters of the dominant (calibration) source, but also discovered a highly correlated low-power multipath, which appears to be due to the reflection of the dominant source from a side-plate of the QFF.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation Grants ECS-0097636 and CCR-0104887, the Swedish Science Council (VR), and the Swedish Foundation for International Cooperation in Research and Higher Education (STINT). The Small Aperture Directional Array datasets used in the generation of results shown in this work were provided under the NASA Langley Research Center Grant No. NAG-1-2133.

Arnold, D. P. (**2001**). *A MEMS-Based Directional Acoustic Array for Aeroacoustic Measurements*, M.S thesis, University of Florida, Gainesville, FL.

Bent, P. H., Guo, Y., Horne, W. C., and Watts, M. E. (**1996**). ''Airframe noise scaling and source localization,'' Proceedings of National Conference on Noise Control Engineering, Vol. 1, 145–150.

Brooks, T. F., and Humphreys, Jr., W. M. (**1999**). ''Effect of directional array size on the measurement of airframe noise components,'' AIAA Paper 99-1958, 5th AIAA/CEAS Aeroacoustics Conference and Exhibit, Bellevue, WA.

Brooks, T. F., Marcolini, M. A., and Pope, D. S. (**1987**). ''A directional array approach for the measurement of rotor noise source distributions with controlled spatial resolution,'' J. Sound Vib. **112**, 192–197.

Dougherty, R. P., and Stoker, R. W. (**1998**). ''Sidelobe suppression for phased array aeroacoustic measurements,'' AIAA Paper 98-2242, 4th AIAA/CEAS Aeroacoustics Conference, 19th AIAA Aeroacoustics Conference, Toulouse, France.

Gramann, R. A., and Mocio, J. (**1993**). ''Aero-acoustic measurements in wind tunnels using conventional and adaptive beamforming methods,'' AIAA Paper 93-4341.

Högbom, J. A. (**1974**). ''Aperture synthesis with a non-regular distribution of interferometer baselines,'' Astron. Astrophys., Suppl. Ser. **15**, 417–426.

Humphreys, Jr., W. M., Brooks, T. F., Hunter, Jr., W. W., and Meadows, K. R. (**1998**). ''Design and use of microphone directional arrays for aeroacoustic measurements,'' AIAA Paper 98-0471, AIAA, 36th Aerospace Sciences Meeting and Exhibit, Reno, NV.

Li, J., and Stoica, P. (**1996**). ''Efficient mixed-spectrum estimation with applications to target feature extraction,'' IEEE Trans. Signal Process. **44**, 281–295.

Li, J., Stoica, P., and Zheng, D. (**1997**). ''Angle and waveform estimation via RELAX,'' IEEE Trans. Aerosp. Electron. Syst. **33**, 1077–1087.

Meadows, K. R., Brooks, T. F., Humphreys, Jr., W. M., Hunter, W. W., and Gerhold, C. H. (**1997**). ''Aeroacoustic measurements of a wing-flap configuration,'' AIAA Paper 97-1595, 3rd AIAA/CEAS Aeroacoustics Conference, Atlanta, GA.

Mosher, M. (**1996**). ''Phased arrays for aeroacoustic testing: Theoretical development,'' AIAA Paper 96-1713, 2nd AIAA/CEAS Aeroacoustics Conference, State College, PA.

Mosher, M., Watts, M. E., Barnes, M., and Bardina, J. (**1999**). ''Microphone array phased processing system (mapps)—phased array system for acoustic measurements in a wind tunnel,'' AIAA Paper 99-5576, 18th ASME Wind Energy Symposium, AIAA, 37th Aerospace Sciences Meeting and Exhibit, Reno, NV.

Mosher, M., Watts, M. E., Jaeger, S. M., and Jovic, S. (**1997**). ''Calibration of microphone arrays for phased array processing,'' AIAA Paper 97-1678, 3rd AIAA/CEAS Aeroacoustics Conference, Atlanta, GA.

Nelder, J. A., and Mead, R. (**1965**). ''A simplex method for function minimization,'' Comput. J. **7**, 308–313.

Piet, J. F., and Elias, G. (**1997**). ''Airframe noise source localization using a microphone array,'' AIAA Paper 97-1643, 3rd AIAA/CEAS Aeroacoustics Conference, Atlanta, GA.

Piet, J. F., Elias, G., and Lebigot, P. (**1999**). ''Localization of acoustic source from a landing aircraft with a microphone array,'' AIAA Paper 99-1811, AIAA Space Technology Conference and Exposition, Albuquerque, NM.

Powell, M. J. D. (**1964**). ''An efficient method for finding the minimum of a function of several variables without calculating derivatives,'' Comput. J. (UK) **7**, 155–162.

Rosenbrock, H. H. (**1960**). ''An automatic method for finding the greatest or least value of a function,'' Comput. J. (UK) **3**, 175–184.

Stoica, P., and Nehorai, A. (**1989**). ''Statistical analysis of two nonlinear least-squares estimators of sine-wave parameters in the colored-noise case,'' Circuits Syst. Signal Process. **8**, 3–15.

Stoica, P., Jakobsson, A., and Li, J. (**1997**). ''Sinusoid parameter estimation in the colored noise case: asymptotic Cramér-Rao bound, maximum likelihood and nonlinear least-squares,'' IEEE Trans. Signal Process. **45**, 2048–2059.

Storms, B. L., Hayes, J. A., Ross, J. C., and Moriarty, P. J. (**1999**). ''Aeroacoustic measurements of slat noise on a three-dimensional high-lift system,'' AIAA Paper 99-1957, 33rd AIAA Thermophysics Conference, Norfolk, VA.

Van Trees, H. L. (**2002**). *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory* (Wiley, New York).

Watts, M. E., Mosher, M., and Barnes, M. J. (**1996**). ''The microphone array phased processing system (mapps),'' AIAA Paper 96-1714, 2nd AIAA/CEAS Aeroacoustics Conference, State College, PA.

Zatman, M. (**1998**). ''How narrow is narrowband?'' IEE Proc. Radar, Sonar Navigat. **145**, 85–91.

Real time inverse filter focusing through iterative time reversal

Gabriel Montaldo,[a] Mickaël Tanter, and Mathias Fink
Laboratoire Ondes et Acoustique, Université Paris VII/ESPCI, CNRS UMR 7587, 10 Rue vouquelin, 75231 Paris Cedex 05, France

(Received 19 July 2002; accepted for publication 30 October 2003)

In order to achieve an optimal focusing through heterogeneous media we need to build the inverse filter of the propagation operator. Time reversal is an easy and robust way to achieve such an inverse filter in nondissipative media. However, as soon as losses appear in the medium, time reversal is not equivalent to the inverse filter anymore. Consequently, it does not produce the optimal focusing and beam degradations may appear. In such cases, we showed in previous works that the optimal focusing can be recovered by using the so-called spatiotemporal inverse filter technique. This process requires the presence of a complete set of receivers inside the medium. It allows one to reach the optimal focusing even in extreme situations such as ultrasonic focusing through human skull or audible sound focusing in strongly reverberant rooms. But, this technique is time consuming and implied fastidious numerical calculations. In this paper we propose a new way to process this inverse filter focusing technique in real time and without any calculation. The new process is based on iterative time reversal process. Contrary to the classical inverse filter technique, this iteration does not require any computation and achieves the inverse filter in an experimental way using wave propagation instead of computational power. The convergence from time reversal to inverse filter during the iterative process is theoretically explained. Finally, the feasibility of this iterative technique is experimentally demonstrated for ultrasound applications. [DOI: 10.1121/1.1636462]

PACS numbers: 43.60.Pt, 43.40.Fz, 43.60.Cg [RLW] Pages: 768–775

I. INTRODUCTION

The basic process in most acoustic applications is to focus a sound or ultrasound beam through an unknown medium. In our laboratory, we have demonstrated that the time-reversal method is very robust and a simple way to focus through complex media such as waveguides, multiple scattering, and random media.[1–4] This focusing technique concerns a wide range of application areas such as ultrasound based medical imaging and therapy,[5–8] underwater acoustics,[9,10] nondestructive testing,[11] room acoustics,[12] and telecommunications.[13,14] The method is based on the time-reversal invariance of the wave equation in a nondissipative medium. A wave diverging from a punctual source and received by a set of transducers enclosing the medium can be time reversed and reemitted by the elements. The generated wave will converge back to its original source location as if time was running backwards. In this case, we can imagine that the time reversal process acts as an inverse filter of the diffraction process.

Unfortunately, there are two problems that degrade the time-reversal focusing and do not permit one to achieve an inverse filter focusing. First, dissipation breaks the time-reversal invariance of the wave equation. A wave will suffer twice the dissipation effects during its forward and backward propagation and this irreversibility degrades the time-reversal focusing quality. As an example, it was shown in the topic of ultrasonic brain imaging and therapy that these dissipation effects may strongly degrade the focusing through the human skull.[15] Second, a time-reversal close cavity[16] is difficult to build and in practice, the time-reversal operation is achieved on a limited aperture known as a time-reversal mirror. This angular aperture limitation gives rise to new information losses that may cause the focusing quality to decrease, particularly in strongly diffracting media.[17]

In order to correct these problems, it is possible to calculate an *explicit* spatiotemporal inverse filter of the propagation by embedding a set of receivers (control points) inside the propagation medium and acquiring experimentally all the impulse responses relating each element of the array to each control point.[18] After the experimental acquisition of this propagation operator we can compute a numerical inversion and deduce the set of optimal focusing signals. In a strongly dissipative medium the results given by this technique are better than the results achieved with a time-reversal process,[19] but the time needed to acquire the operator and compute its numerical inversion is very long. Moreover, a regularization process must be applied to the operator in order to clean noise off before inversion and this regularization step requires some *a priori* information input in the method. In this article we propose a new method which allows one to recover the inverse filter focusing using an iterative time-reversal process. This iterative process may be implemented in real time as it introduces very simple operations and does not require one to acquire all the propagation operators relating the array to the control points. Here, the numerical inversion is replaced by an ''experimental inversion'' achieved by successive wavefield propagation in the medium. Finally, from a theoretical point of view, it is explained here how time reversal is associated to inverse filter and how iterative time-reversal operations are capable of improving

[a]Electronic mail: gabrielmontaldo@yahoo.fr

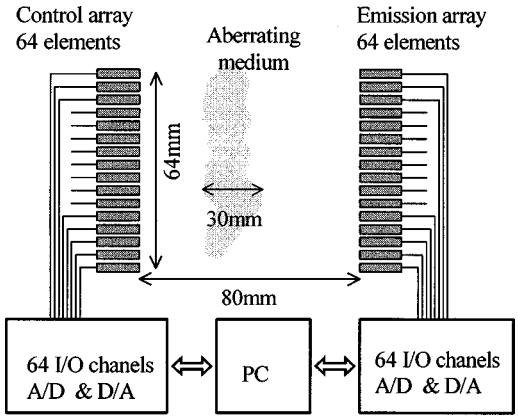


FIG. 1. Experimental setup.

step by step the focusing and eventually reach the optimal inverse filter focusing both in space and time.

II. DESCRIPTION OF THE METHOD

The experimental configuration consists of an *emitting* array of transducers and a *control* array inside the propagation medium. Both arrays (*emitting* and *control*) consist of transducers which are able to work in transmitting and receiving mode. The medium is considered as heterogeneous and dissipative (see Fig. 1).

In the following discussions, one should notice that the notation $s(t)$ describes a vector of signals $s(t)=[s_1(t),s_2(t),\ldots,s_i(t),\ldots,s_N(t)]$ received or transmitted by the elements of one array. Our objective is to focus a sharp pulse over the mth transducer of the control array and 0 over all the other transducers. We can write this objective like a vector of signals $o(t)=[0,\ldots,0,o_m(t),0,\ldots,0]$, where $o_m(t)$ is a sharp signal. The focusing of this signal is done in the following sequence.

(1) The first step of the process consists in transmitting the time reversed signal $o(-t)$ from the *control* array. After propagation through the medium, the resulting wave front $e_1(t)$ is recorded on the *emitting* array [see Fig. 2(a)].

(2) In a second step, these signals are time reversed and $e_1(-t)$ is retransmitted by the *emitting* array. As one can notice, the description below corresponds to a classical time-reversal experiment. However, in a dissipative medium, time reversal is not optimal and the focusing pattern $r_1(t)$ recorded on the *control* array is not exactly the objective $o(t)$ and it has large sidelobes [Fig. 2(b)].

(3) In order to recover an optimal focusing, a third step is introduced. The difference between the focusing pattern $r_1(t)$ and the desired spatiotemporal objective $o(t)$ is then computed. This difference $d(t)=o_1(t)-r(t)$ is time reversed and transmitted by the *control* array. After propagation through the medium, a wave $c(t)$ is received and recorded on the *emitting* array. If we retransmit this signal $c(-t)$ by the *emitting* array, we should obtain on the control array a quite good approximation of the sidelobes $d(t)$.

(4) $c(t)$ can be considered as a correction signal to improve the focusing pattern. Thus, by transmitting the corrected set of signals $e_2(t)=e_1(t)+c(t)$ on the *emitting* array, we will obtain in the *control* array a new focusing $r_2(t)$

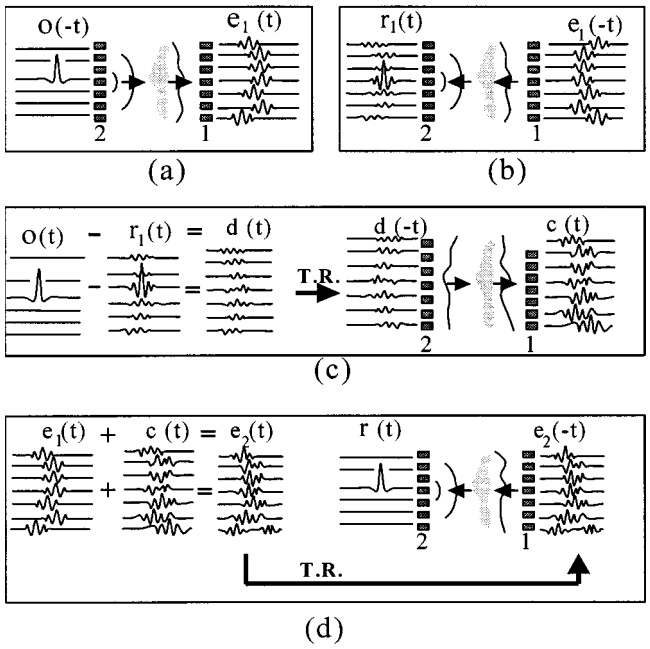


FIG. 2. The iterative method. (a) Emission of the objective pattern, (b) focusing by time-reversal process, (c) emission of the difference wave field, (d) focusing with new the corrected signals.

that approaches in a better way the focusing objective. This process can be repeated iteratively up to the attainment of the optimal focusing.

In Fig. 3 we can see an experimental focusing through an aberrating and dissipative medium using time reversal and using this iterative process. After the iteration, the spatiotemporal lobes decrease in roughly 20 dB.

Due to the simplicity of all these operations (emission, reception, time reversal, signal subtraction), this iteration can be implemented with a very high speed allowing a real time correction of the aberrations of the medium while focusing on each control point. Moreover, contrary to the classical explicit spatiotemporal inverse filter technique, it is not necessary to acquire all the impulse responses relating each element of the array to each control point. Finally, from a

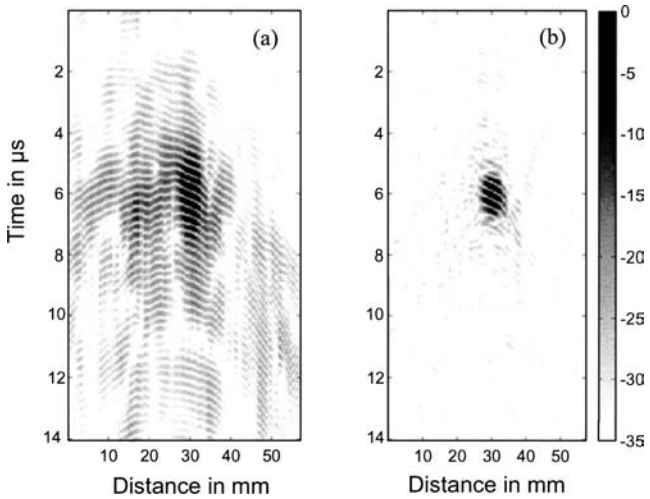


FIG. 3. Spatiotemporal representation of the focusing achieved through an aberrating mask. (a) with time-reversal, (b) after 20 iterations of correction.

theoretical point of view, it is interesting to note that this iterative time-reversal process allows us to build a propagation inverse filter, as will be shown in Sec. III.

III. THEORETICAL BACKGROUND

We consider a system composed by N_e *emitting* transducers and N_c *control* transducers. The propagation between the *emitting* array and the *control* array can be described by the set of impulse responses $h_{ij}(t)$, $i=1,...,N_c$, $j=1,...,N_e$ relating the *emitting* transducer *j* and the *control* transducer *i*. This set of impulse responses is called the propagation operator.

If we send a signal $e_j(t)$, $j=1,...,N_e$ from the *emitting* array we can calculate the signal $r_i(t)$, $i=1,...,N_c$ obtained in the *control* array as

$$r_i(t)=\sum_j h_{ij}(t)\otimes e_j(t). \tag{1}$$

Or, in the Fourier domain

$$R_i(\omega)=\sum_j H_{ij}(\omega)E_j(\omega). \tag{2}$$

That may be written in matrix form as

$$R(\omega)=\mathbf{H}(\omega)E(\omega). \tag{3}$$

Due to the spatial reciprocity, if we send a signal $c_i(t)$, $i=1,...,N_c$ from the *control* array we will obtain a signal $d_j(t)$, $j=1,...,N_e$ in the *emitting* array related by the transposed matrix ${}^t\mathbf{H}$,

$$D(\omega)={}^t\mathbf{H}(\omega)C(\omega), \tag{4}$$

then $\mathbf{H}$ permits one to calculate the propagation from the *emitting* to the *control* array and ${}^t\mathbf{H}$ the propagation from the *control* to the *emitting* array.

Our objective is to focus a short pulse in the *m*th element of the *control* array that we can write as $o(t)=[0,...,\delta(t),\ldots,0]$ or in the Fourier domain $O(\omega)=[0,...,1,...,0]$.

A time reversal focusing begins by sending $o(-t)$ from the *control* array. $o(-t)$ is in the Fourier domain $O^*(\omega)$ where the asterisk notates the complex conjugation, then we can calculate the signals received on the *emitting* array after the propagation as

$$E(\omega)={}^t\mathbf{H}(\omega)O^*(\omega). \tag{5}$$

We time reversed this signal $E^*(\omega)$ and we sent it back to the *control* array, the signal $F(\omega)$ received after the propagation is

$$F(\omega)=H(\omega)E^*(\omega)=\mathbf{H}(\omega)\mathbf{H}^\dagger(\omega)O(\omega), \tag{6}$$

where † denotes the transpose conjugate: $\mathbf{H}^\dagger={}^t\mathbf{H}^*$. Thus, the focusing quality obtained by time-reversal processing is directly linked to the operator $\Delta(\omega)=\mathbf{H}(\omega)\mathbf{H}^\dagger(\omega)$ called *time-reversal operator*. If the system consists in a nondissipative close cavity and if the set of transmitters fully map the boundary surface of the cavity, a time-reversal operation will tend to recreate the inverse propagation operator of the system

$$\begin{aligned}&\Delta=\mathbf{H}(\omega)\mathbf{H}^\dagger(\omega)\cong\mathbf{I}(\omega),\\ &\mathbf{H}^{-1}(\omega)\cong\mathbf{H}^\dagger(\omega),\\ &h_{ji}^{-1}(t)\cong h_{ij}(-t).\end{aligned} \tag{7}$$

Nevertheless, as soon as information losses appear during a time-reversal experiment, for example, in a open or a lossy system, time reversal is not able to fully recover the inverse of the propagation anymore. In such situations, it is not ensured that time reversal will be able to achieve a good focusing quality. The technique that we propose tries to iterate experimentally time-reversal operations in order to improve the focusing result and finally recover the optimal focusing achieved by inverse filtering.

The first step of this method is the standard time-reversal focusing described before. We choose an objective $O(\omega)$, this objective is time reversed $O^*(\omega)$ and sent from the *control* array. After propagation, the wave form received on the *emitting* array is

$$E_1(\omega)={}^t\mathbf{H}(\omega)O^*(\omega). \tag{8}$$

These signals are time reversed and sent from the *emitting* array to the *control* array. The wave field induced on the *control* array can be written as

$$R_1(\omega)=\mathbf{H}(\omega)E_1^*(\omega). \tag{9}$$

The difference between this focusing result and the desired focusing objective is

$$D_1(\omega)=O(\omega)-R_1(\omega). \tag{10}$$

This wave field difference is then time reversed and $D_1^*(\omega)$ is transmitted by the *control* array. The resulting wave field propagates through the medium and a signal ${}^t\mathbf{H}(\omega)D_1^*(\omega)$ is recorded on the *emitting* array. This signal can be seen as a correcting wave field that can be added after time reversal to the initial emission signals $E_1(\omega)$. We built a new emission signal

$$E_2(\omega)=E_1(\omega)+{}^t\mathbf{H}(\omega)D_1^*(\omega). \tag{11}$$

After time-reversal emission and propagation, the new focusing result on the *control* array will be

$$R_2(\omega)=\mathbf{H}(\omega)E_2^*(\omega). \tag{12}$$

This iterative process can continue in order to correct step by step the emission signals $E_n(\omega)$,

$$\begin{aligned}&R_n=\mathbf{H}E_n^*,\\ &D_n=O-R_n,\\ &E_{n+1}=E_n+{}^t\mathbf{H}D_n^*\end{aligned} \tag{13}$$

with an initial step $E_1(\omega)={}^t\mathbf{H}(\omega)O^*(\omega)$. At this point, the basic question is whether this iteration converges to the solution. Note that for the sake of clarity, we forget the ω dependence in the following equations. As we are only interested in the result of R_n, we can write the iteration process (13) as a function of $R_n(\omega)$,

$$R_{n+1}=R_n-\mathbf{\Delta}(R_n-O). \tag{14}$$

It is easy to demonstrate that the *n*th iteration of Eq. (14) is

$$R_n = [\mathbf{I} - (\mathbf{I} - \mathbf{\Delta})^n] O, \tag{15}$$

where $\mathbf{I}$ is the identity matrix. The difference between the focusing R_n at the nth iteration and the initial focusing objective O is

$$D_n = (\mathbf{I} - \mathbf{\Delta})^n O. \tag{16}$$

This difference must converge to 0 in an exact focusing. In order to analyze the convergence of D_n, we consider the diagonalization of $\mathbf{\Delta}$,

$$\mathbf{\Delta} = \mathbf{P}\mathbf{S}\mathbf{P}^{\dagger}, \tag{17}$$

where $\mathbf{S}$ is a real and diagonal matrix containing the eigenvalues λ_i, $i = 1,...,N_c$ of the time-reversal operator in a decreasing order, and $\mathbf{P}$ is a unitary matrix. The set of eigenvalues can be divided into two parts: the M non-null eigenvalues, which describe the physically relevant data contained in the propagation operator, and the null ones. As shown in Ref. 18, the number of physically relevant eigenvalues of $\mathbf{H}\mathbf{H}^{\dagger}$ is linked to the number of independent plane wave components that can propagate from the *emitting* array to the *control* array. Thus, this number M will depend on all the parameters defining the experimental setup: the number of emitting and control transducers, the aperture of both arrays, the distance from one array to the other, the central emission frequency and bandwidth, and the heterogeneities of the propagating medium. As an example, if the number of elements of the *control* array N_c is greater than the number of elements of the *emitting* array N_e, the rank of the null eigenvalues subspace will be at least the difference $N_e - N_c$. This number of null eigenvalues can also increase when two or more elements are very close. In that particular situation the impulse responses of two neighbors are strongly correlated and similar columns appear in the operator $\mathbf{H}$ so generating null eigenvalues in $\mathbf{\Delta}$.

How do null and non-null eigenvalues behave, respectively, during the iterative process? According to Eqs. (16) and (17), the nth difference will be

$$D_n = \mathbf{P}(\mathbf{I} - \mathbf{S})^n \mathbf{P}^{\dagger} \cdot O. \tag{18}$$

The eigenvalues of $(\mathbf{I} - \mathbf{S})^n$ are $(1 - \lambda_i)^n$, and they converge to 0 if $0 < \lambda_i < 2$.

Thus, this difference between the focusing objective and the focusing pattern obtained at iteration n will tend to zero if all the eigenvalues are understood between 0 and 2. As $\mathbf{\Delta}$ is Hermitian and positive definited, its eigenvalues are real and positive. Besides, by normalizing adequately the propagation operator $\mathbf{H}$ we can easily force all the eigenvalues to be between 0 and 2. (At each iteration we record, amplify, and re-send the signals by changing this amplification factor we can adjust the normalization of the propagation matrix.) In other words, if the eigenvalues are different from zero, the difference D_n will tend to zero during the iterative process. On the contrary, the null eigenvalues will remain during the iterative process. Thus, we can decompose the eigenvalues space into two distinct subspaces corresponding to the *non-null eigenvalues subspace* and the *null eigenvalues subspace*. So, the iterative process allows one to project the focusing objective O over the *non-null subspace*. It will remain a residual D_∞ corresponding to the projection of the objective over the *null subspace*

$$D_\infty = \mathbf{P} \begin{bmatrix} 0 & & & & & \\ & \ddots & & & 0 & \\ & & 0 & & & \\ & & & 1 & & \\ & 0 & & & \ddots & \\ & & & & & 1 \end{bmatrix} \mathbf{P}^{\dagger} \cdot O. \tag{19}$$

The speed of convergence depends on the eigenvalues λ_i, when an eigenvalue is near 0 $(1 - \lambda_i)^n$ converges slowly to 0, there is a compromise between the time of convergence and the focusing quality.

A. Time reversal, the inverse filter, and the iterative method

In order to compare the focusing obtained by time reversal and inverse filter we are going to introduce the singular values decomposition of the propagation matrix $\mathbf{H}$,

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{P}^{\dagger}, \tag{20}$$

where $\mathbf{D}$ is an $N_c \times N_c$ diagonal real matrix with the singular values α_i in the diagonal, $\mathbf{U}$ and $\mathbf{P}$ are, respectively, $N_c \times N_c$ and $N_c \times N_e$ unitary matrices. The time reversal operator can be written as

$$\mathbf{\Delta} = \mathbf{P}\mathbf{D}^2\mathbf{P}^{\dagger}. \tag{21}$$

Here one can notice that the eigenvalues λ_i $i = 1,...,N_c$ of the time-reversal operator are the squares of the singular values of $\mathbf{H}$, $\lambda_i = \alpha_i^2$.

On the contrary, the inverse filter technique proposes an explicit inversion of $\mathbf{H}$ based on the singular value decomposition. This inversion is only achieved with the M non-null singular vectors of the decomposition leading to an approximation $\hat{\mathbf{H}}^{-1}$ of the inverse of $\mathbf{H}$,

$$\hat{\mathbf{H}}^{-1} = \mathbf{P}\hat{\mathbf{D}}^{-1}\mathbf{U}^{\dagger} = \mathbf{P} \begin{bmatrix} \alpha_1^{-1} & & & & \\ & \ddots & & 0 & \\ & & \alpha_M^{-1} & & \\ & 0 & & 0 & \\ & & & & 0 \end{bmatrix} \mathbf{U}^{\dagger}. \tag{22}$$

Therefore, the eigenvalues of $\mathbf{H}\hat{\mathbf{H}}^{-1}$ are equal to 1 or 0.

The iterative method is described by the operator $[\mathbf{I} - (\mathbf{I} - \mathbf{\Delta})^n]$. Thus, the eigenvalues of this operator after n iterations are equal to $[1 - (1 - \lambda_i)^n]$. In this case, the eigenvalues will converge potentially to 1, as $0 < \lambda_i < 2$.

The iterative time-reversal process tends to recreate the propagation inverse filter on the subspace of the highest singular values of the propagation operator $\mathbf{H}$. The results obtained by using these three different methods are described in terms of eigenvalues and summarized in Table I.

TABLE I. Operators and eigenvalues of the three focusing techniques.

	Operator	Eigenvalue	Error
Inverse filter	$\mathbf{H}\hat{\mathbf{H}}^{-1}$	1 for highest, 0 for lowest	$\sigma_{if}^2=\sigma^2\sum_{i=1}^{M}\frac{\lvert b_i\rvert^2}{\lambda_i}+N_c\sigma^2$
Time reversal	$\mathbf{\Delta}=\mathbf{H}\mathbf{H}^{\dagger}$	λ_i	$\sigma_{\mathrm{tr}}^2=2N_c\sigma^2$
Iterative inv. filt.	$\mathbf{I}-(\mathbf{I}-\mathbf{\Delta})^n$	$1-(1-\lambda_i)^n$	$\sigma_{\mathrm{it}}^2=3N_c\sigma^2$

B. Noise influence

The above-mentioned formalism was not considering the random noise disturbing the acquisition system during the experiment. The Appendix deals with the exact calculation of the mean quadratic error induced by the three methods and we report here the main results.

For the classical spatiotemporal inverse filter, the propagation matrix $\tilde{\mathbf{H}}$ measured at each frequency is the sum of the true propagation matrix $\mathbf{H}$ and a noise matrix $\mathbf{A}$, i.e., $\tilde{\mathbf{H}}=\mathbf{H}+\mathbf{A}$. Due to this random noise, the null singular values of the matrix $\mathbf{H}$ are replaced in $\tilde{\mathbf{H}}$ by low singular values without physical relevance. However, during inversion the magnitudes of $1/\alpha_i$ will become very large and will generate a noise amplification. We can define a quadratic deviation of the inverse filter focusing as $\sigma_{if}^2=\langle\lvert R-\langle R\rangle\rvert^2\rangle$, where R is the final focusing and $\langle\ \rangle$ is the average over noise realizations.

If we assume that all the components A_{ij} of the noise matrix are independent and have the same quadratic deviation σ, the quadratic error of the focusing pattern achieved by the inverse filter is

$$\sigma_{if}^2=\sigma^2\sum_{i=1}^{M}\frac{\lvert b_i\rvert^2}{\lambda_i}+N_c\sigma^2, \tag{23}$$

where b_i are the coefficients of the decomposition of the objective vector O over the M highest eigenvalues V_i of the time reversal operator $\mathbf{\Delta}$, i.e., $b_i=V_iO$.

As one can notice in Eq. (23), if the lowest eigenvalues are kept during the inversion process, the noise will be highly amplified. For this reason, we need to have *a priori* knowledge of the noise level in order to substitute all the singular values which are lower than this noise level by 0 in the inverse matrix [see Eq. (22)].

On the contrary, in the case of the iterative process, a new independent noise is added at each step of the iteration and the final quadratic error is

$$\sigma_{it}^2=3N_c\sigma^2. \tag{24}$$

It is also the case for the two-step process achieved by the time-reversal technique as shown in the Appendix the time-reversal focusing noise is

$$\sigma_{\mathrm{tr}}^2=2N_c\sigma^2. \tag{25}$$

The mean quadratic errors of the time-reversal and the iterative inverse filter processes are independent of the objective and are very similar in magnitude. This result explains the fact that these processes are very robust and noise independent in opposition to the classical inverse filter technique that is based on an explicit inversion of the propagation operator. Table I summarizes the comparison of the three methods in terms of operators, eigenvalues, and noise.

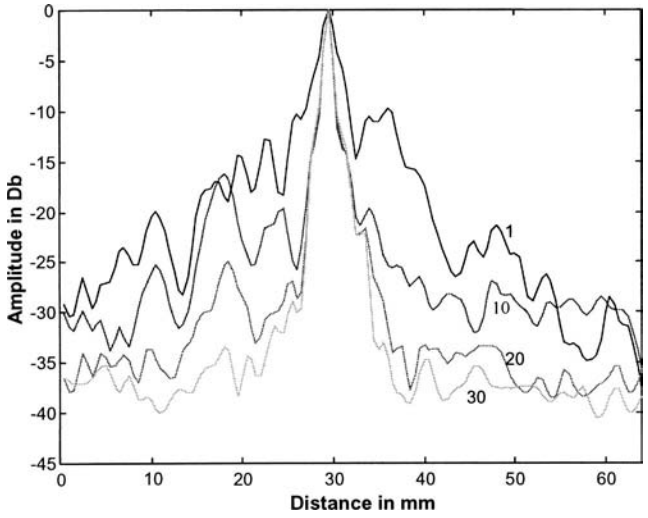


FIG. 4. Spatial distribution of the focal spot with 1, 10, 20, and 30 iterations.

IV. EXPERIMENTAL RESULTS

In order to test experimentally the potential of the iterative inverse filter method, we use a multichannel system made of two arrays of 64 elements working at a central frequency $f_0=1.5$ MHz. Two arrays are located on both sides of a highly aberrating and lossy medium consisting of an irregular block of polymer of approximately 3 cm width, the mean attenuation of the block is 15 dB with spatial variations of nearly 10 dB between different points. The maximal phase differences induced by the block in different points are nearly $5\times\pi$ (Fig. 1). Each element is linked to its own fully programmable Emission/Reception electronic working at a 50 MHz sampling rate. The length of the recorded signals was 20 μs in this experiment.

The focusing objective $O_i(t)$ consists in a Gaussian-shaped signal on one transducer of the *control* array,

$$O_{i_0}(t)=e^{-(t-T)^2/\sigma^2}\sin(2\pi f_0 t),$$
$$O_i(t)=0,\quad i\neq i_0. \tag{26}$$

Figure 4 shows the evolution of the measured focusing pattern in the *control* array for different number of iterations. The case of a single iteration is the result of the standard time-reversal technique. Due to the strong aberration of the medium the sidelobes are very high, nearly 10 dB. After some iterations these sidelobes are cleared off and here 30 iterations are enough to reach a saturation in the focusing quality. The improvement of the focusing pattern between the first and the last iteration is extremely important. The resolution is improved by a factor 2 and the sidelobe level is reduced in nearly 15 dB.

On a temporal point of view, the signal received at the focus (on the i_0th transducer of the *control* array) is shown in Fig. 5 for the first and the last iteration. The first iteration (corresponding to time-reversal focusing) is a larger and

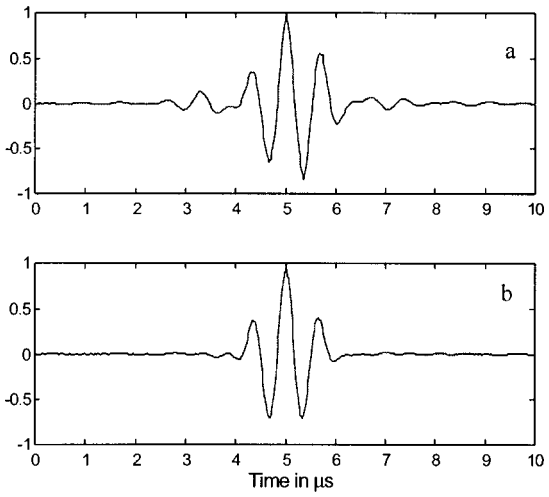


FIG. 5. (a) Wave shape with 1 iteration. (b) Wave shape after 30 iterations.

more distorted wave shape than in the final iteration. The frequency bandwidth of the focal signal received at the last iteration is larger than for the first iteration. This time compression of the signal is due to the frequency bandwidth correction achieved by the iterative process. Step by step, a more important weight is given back to the frequency components at the edges of the bandwidth compared to the central ones. In other words, the iteration does a deconvolution of the transducers' bandwidth.

Figure 6 compares this result with the direct inversion of the propagation matrix and with the classical time-reversal process. The spatiotemporal inverse filter and the iterative method give the same focusing quality. On the other hand, due to the heterogeneous absorption induced during the propagation through the aberrating medium, the time-reversal process does not give good results. The difficulty in focusing through such an aberrating medium is demonstrated in the same figure by the focusing pattern achieved with a classical cylindrical time delay law (i.e., assuming a constant speed of sound in the medium). As one can notice, the cylindrical law is not able to focus on the correct point and the

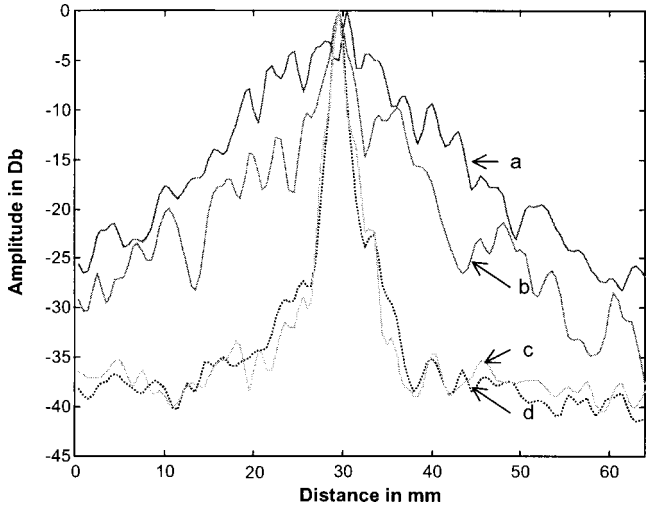


FIG. 6. Focusing patterns obtained by (a) classical cylindrical time delay focusing law, (b) time-reversal processing, (c) iterative time reversal processing, (d) spatiotemporal inverse filter.

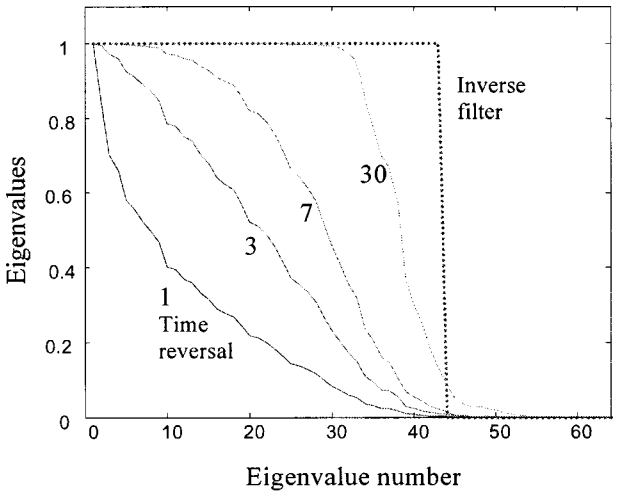


FIG. 7. Singular value distribution of the operator $[\mathbf{I}-(\mathbf{I}-\mathbf{\Delta})^n]$ after $n=1$, $n=3$, $n=7$, $n=30$ iterations, and the optimal singular value distribution of $\mathbf{H}\hat{\mathbf{H}}^{-1}$ obtained by the inverse filter.

sidelobes level is very high, of about -5 dB.

Finally, one should notice that the singular value distribution of the operator achieved by the iterative process evolves toward the inverse filter distribution of $\mathbf{H}\hat{\mathbf{H}}^{-1}$. All the singular values tend to be equalized during the iterative proces (see Fig. 7).

It is interesting to compare the total time involved in a complete focusing over the 64 elements with an iterative technique and with a numerical inverse filter. The iterative technique is limited by the travel time of the waves in the medium and we can estimate this time as $T_p \times N_c \times N_{\mathrm{it}}$, where T_p is the time of propagation, N_c is the number of points to focus, and N_{it} is the number of iterations.

In the case of the numerical inversion we need to acquire the propagation matrix that takes a time $T_p \times N_c$, the numerical inversion involves $N_c \times N_e$ Fourier transforms, and a SVD decomposition for each frequency. The Fourier transform is an algorithm of $n_t \times \log_2(n_t)$ order where n_t is the number of samples of the signal. The SVD is an algorithm of $N_c \times N_e$ order, as the number of frequencies is of n_t order, the total algorithm is of $N_c \times N_e \times n_t \times \log_2(n_t)$ order. Of course the exact comparison of time depends on the computer where we implement the calculations. In the example we implemented the iterative process learns to focus over the 64 points in 0.3 s and a complete numerical decomposition takes 6.2 s. Optionally, if we are interested in focusing only in a given location the iterative process takes only 4.7 ms.

V. CONCLUSION

A new real time method for optimal focusing through absorbing and complex media was presented. This method iterates the time-reversal process in order to cancel the focusing beam degradation due to absorption and other information losses. Contrary to numerical inverse filtering, this iterative technique does not require any computation and it achieves the inverse filter in an experimental way using wave propagation instead of computer power. This iteration may be implemented at real time and it has a lot of potential applications in imaging and communications. By analyzing

the eigenvalues of the time-reversal operator we have demonstrated that the iteration of the time-reversal process converges to the inverse filter. Experimentally, we have proved the efficiency of the iteration in an absorbing and aberrating medium and its results are very close to those of the inverse filter. We think that this method may be very useful to focus in complex media that change slowly in time. In this case, the inverse filter is not capable of focusing because while the whole propagation operator is recorded and inverted the medium changes. As the time of one iteration is of the order of the propagation time, we can correct the focusing continuously with a few iterations starting each correction with the last focusing. This iterative method is particularly suited to the case of short propagation times. In such configurations the time need by this technique is shorter than in the case of a numerical inversion.

APPENDIX

The aim here is to demonstrate that when the experimental noise is considered, the iterative inverse filter process is more stable than the classical spatiotemporal inverse filter based on the explicit inversion of the propagation operator. In order to compare both techniques, we will introduce the mean quadratic error $\sigma^2_{\text{focusing}}$ corresponding to the mean quadratic deviation of the focusing pattern finally obtained R as

$$\sigma^2_{\text{focusing}} = \langle |R - \langle R \rangle|^2 \rangle. \tag{A1}$$

where $\langle\ \rangle$ denotes an average over noise realizations.

The numerical inverse filter needs the measurement of the propagation matrix, this measurement suffers a random noise and then we have access to an experimental matrix,

$$\tilde{\mathbf{H}} = \mathbf{H} + \mathbf{A}, \tag{A2}$$

where $\tilde{\mathbf{H}}$ is the measured matrix, $\mathbf{H}$ is the true propagation matrix, and $\mathbf{A}$ is a matrix of noise. If the noise level is not very high we are able to approximate the inverse $\tilde{\mathbf{H}}^{-1}$ by a first-order development

$$\tilde{\mathbf{H}}^{-1} = \mathbf{H}^{-1}(\mathbf{I} - \mathbf{H}^{-1}\mathbf{A}). \tag{A3}$$

The emission signal corresponding to a focusing objective O is equal to $E = \tilde{\mathbf{H}}^{-1} O$ and the focusing pattern R achieved on the *control* array while emitting E on the *emitting* array is

$$R = \mathbf{H}E + a = O - \mathbf{H}^{-1}\mathbf{A}O + a, \tag{A4}$$

where a is a vector corresponding to a new random realization of noise during the experimental reception step. Using Eq. (A1), the random error due to this method can be calculated as the mean quadratic deviation of the two last terms. The noise $\mathbf{A}$ is independent of a and the quadratic deviation is

$$\sigma^2_{\text{FI}} = \langle |\mathbf{H}^{-1}\mathbf{A}O|^2 \rangle + \langle |a|^2 \rangle. \tag{A5}$$

Assuming that all the elements a_i of the noise vector are independent and have the same quadratic deviation $\langle |a_i|^2 \rangle = \sigma^2$ it is possible to write the last term as $\langle |a|^2 \rangle = N_c \sigma^2$. In order to calculate the first term, the index notation is introduced

$$\begin{aligned} \langle |\mathbf{H}^{-1}\mathbf{A}O|^2 \rangle &= \langle O_i^* A_{ij}^* H_{jk}^{-1^*} H_{kl}^{-1} A_{lm} O_m \rangle \\ &= O_i^* \Delta_{jl}^{-1} O_m^* \langle A_{ij}^* A_{lm} \rangle. \end{aligned} \tag{A6}$$

Considering that all the noise components A_{ij} are independent of each other and that they have the same quadratic deviation then $\langle A_{ij}^* A_{lm} \rangle = \sigma^2 \delta_{il} \delta_{jm}$. We can conclude that

$$\langle |\mathbf{H}^{-1}\mathbf{A}O|^2 \rangle = \sigma^2 O_l^* \Delta_{jl}^{-1} O_j = \sigma^2 O^* \boldsymbol{\Delta}^{-1} O. \tag{A7}$$

The objective vector can be decomposed over the eigenvector base of $\boldsymbol{\Delta}$, $O = \Sigma_{i=1}^{N_c} b_i V_i$ and Eq. (A7) can be rewritten as

$$O^* \boldsymbol{\Delta}^{-1} O = \sum_{i=1}^{N_c} |b_i|^2 / \lambda_i, \tag{A8}$$

where λ_i are the eigenvalues of the operator $\boldsymbol{\Delta}$. Equation (A8) clearly illustrates the fact that the noise is amplified by the lowest eigenvalues during the explicit inversion of the propagation operator. When the eigenvalues become too low, it is necessary to change $1/\lambda_i$ by zero in the inverse matrix.[18] However, this regularization step before inversion requires an *a priori* knowledge of the noise level (i.e., the singular values that must be rejected during the inversion process).

Thus, the mean quadratic error corresponding to the classical inverse filter process is finally

$$\sigma^2_{\text{FI}} = \sigma^2 \sum_{i=1}^{M} \frac{|b_i|^2}{\lambda_i} + N_c \sigma^2, \tag{A9}$$

where M is the number of the highest singular values kept during the explicit inversion process.

On the other hand, how does the noise behaves in the iterative process?

We can see that the noise is incorporated at two stages during the iteration process. First, a noise vector a_n appears during the reception of R_n on the *control* array in Eq. (13). Second, a second noise b_n appears during the backpropagation of the difference. Thus, the iterative process including noise can be described as follows:

$$\begin{aligned} R_n &= \mathbf{H}E_n^* + a_n, \\ D_n &= O - R_n, \\ E_{n+1} &= E_n + \mathbf{H}D_n^* + b_n. \end{aligned} \tag{A10}$$

This iteration can be described only in terms of R,

$$R_{n+1} = R_n - \boldsymbol{\Delta}(R_n - O) + \mathbf{H}b_n^* - a_n + a_{n+1}. \tag{A11}$$

Similar to Eq. (15), an explicit equation for the nth iteration R_n can be given

$$\begin{aligned} R_n = [\mathbf{I} - (\mathbf{I} - \boldsymbol{\Delta})^n] O + \sum_{j=0}^{n-1} (\mathbf{I} - \boldsymbol{\Delta})^j H b_{n-j}^* \\ + \sum_{j=0}^{n-2} (\mathbf{I} - \boldsymbol{\Delta})^j \boldsymbol{\Delta} a_{n-1-j} + a_n. \end{aligned} \tag{A12}$$

The last three terms indicate the error due to the random noise. All the noise vectors a_j and b_j are independent of each other and the mean quadratic deviation is

$$\sigma_{\mathrm{IT}}^2=\sum_{j=0}^{n-1}\langle|(\mathbf{I}-\mathbf{\Delta})^j\mathbf{H}b_{n-j}|^2\rangle + \sum_{j=0}^{n-2}\langle|(\mathbf{I}-\mathbf{\Delta})^j\mathbf{\Delta}a_{n-1-j}|^2\rangle+\langle|a_n|\rangle. \quad \text{(A13)}$$

If we consider for example an element of the second term, the mean quadratic deviation can be calculated by expressing it in index notation

$$\langle|(\mathbf{I}-\mathbf{\Delta})^j\mathbf{\Delta}a|^2\rangle=[(\mathbf{I}-\mathbf{\Delta})^{2j}\mathbf{\Delta}^2]_{ij}\langle a_i^* a_j\rangle. \quad \text{(A14)}$$

where a_i are components of a generic noise vector a. According to the independence of the components a_i and assuming that all the components have the same quadratic deviation $\langle a_i^* a_j\rangle=\delta_{ij}\sigma^2$, we obtain the trace of the matrix that can be expressed as the sum of the eigenvalues

$$\langle|(\mathbf{I}-\mathbf{\Delta})^j\mathbf{\Delta}a|^2\rangle=\sigma^2\,\mathrm{Trace}[(\mathbf{I}-\mathbf{\Delta})^{2j}\mathbf{\Delta}^2] =\sigma^2\sum_{i=1}^{N_c}(1-\lambda_i)^{2j}\lambda_i^2, \quad \text{(A15)}$$

where λ_i are the eigenvalues of the time-reversal operator. By doing the same procedure for the other elements in the second term of Eq. (A9), we arrive at a geometrical series in each eigenvalue and it can be added if we consider a large number of iterations n,

$$\sum_{j=1}^{n=\infty}\langle|(\mathbf{I}-\mathbf{\Delta})^j\mathbf{\Delta}a_{n-1-j}|^2\rangle=\sigma^2\sum_{i=1}^{N_c}\sum_{j=0}^{n=\infty}(1-\lambda_i)^{2j}\lambda_i =\sigma^2\sum_{i=1}^{N_c}\frac{\lambda_i}{2-\lambda_i}. \quad \text{(A16)}$$

The same calculus can be derived for the first term in Eq. (A9) and the total error is written

$$\sigma_{\mathrm{IT}}^2=\sigma^2\sum_{i=1}^{N_c}\frac{1+\lambda_i}{2-\lambda_i}+N_c\sigma^2. \quad \text{(A17)}$$

As the eigenvalues λ_i are comprised between 0 and 1, we can consider that $(1+\lambda)/(2-\lambda)<2$ and the error of the iteration will verify

$$\sigma_{\mathrm{IT}}^2<3N_c\sigma^2. \quad \text{(A18)}$$

If we compare this error with the one obtained by the "explicit" inverse filter technique, we can see that the iterative technique gives a much more stable result as it remains independent of the eigenvalues and the chosen objective.

Finally, the error induced in a time-reversal process corresponds to the error induced during the first iteration of the iterative process [Eq. (A10)]. We can calculate the quadratic deviation as $\sigma_{\mathrm{TR}}^2=\langle|\mathbf{H}b|^2\rangle+\langle|a|^2\rangle$. Performing a similar calculation, it is easy to conclude that the time reversal error is equal to

$$\sigma_{\mathrm{TR}}^2<2N_c\sigma^2. \quad \text{(A19)}$$

[1] M. Fink, "Time reversal of ultrasonic fields. II. Basic Principles," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 554–566 (1992).
[2] M. Fink, "Time-reversed acoustics," Sci. Am. 91–97 (1999).
[3] A. Derode, A. Tourin, and M. Fink, "Time reversal in multiply scattering media," Ultrasonics **36**, 443–447 (1998).
[4] D. R. Jackson and D. R. Dowling, "Phase conjugation in underwater acoustics," J. Acoust. Soc. Am. **89**, 171–181 (1991).
[5] J. L. Thomas, F. Wu, and M. Fink, "Time reversal mirror applied to litotripsy," Ultrason. Imaging **18**, 106–121 (1996).
[6] D.-L. Liu and R. C. Waag, "Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation," J. Acoust. Soc. Am. **96**, 649–660 (1994).
[7] G. Montaldo, P. Roux, A. Derode, C. Negreira, and M. Fink, "Generation of very high pressure pulses with 1-bit time reversal in a solid waveguide," J. Acoust. Soc. Am. **110**, 2849 (2001).
[8] R. Seip, P. VanBaren, and E. Ebbini, "Dynamic focusing in ultrasound hyperthermia using implantable hydrophone arrays," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 706–713 (1994).
[9] H. C. Song, W. A. Kuperman, W. S. Hodgkiss, T. Akal, C. Ferla, and D. R. Jackson, "Iterative time reversal in the ocean," J. Acoust. Soc. Am. **105**, 3176–3184 (1999).
[10] W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, "Phase conjugation in the ocean: Experimental demonstration of a time reversal mirror," J. Acoust. Soc. Am. **103**, 25–40 (1998).
[11] N. Chakroun, M. Fink, and F. Wu, "Time reversal processing in non destructive testing," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 1087–1098 (1995).
[12] O. Kirkeby and P. A. Nelson, "Digital filter design for inversion problems in sound reproduction," J. Audio Eng. Soc. **47**, 583–595 (1999).
[13] P. Blomgren, G. Papanicolaou, and H. Zhao, "Super-resolution in time reversal acoustics," J. Acoust. Soc. Am. **111**, 230–248 (2002).
[14] S. H. Simon, A. L. Moustakas, M. Stoytchev, and H. Safar, "Communication in a disordered world," Phys. Today **54**(9), (2001).
[15] M. Tanter, J.-L. Thomas, and M. Fink, "Focusing and steering through absorbing and aberating layers: Application to ultrasonic propagation through the skull," J. Acoust. Soc. Am. **103**, 2403–2410 (1998).
[16] D. Cassereau and M. Fink, "Time reversal of ultrasonic fields. III. Theory of the closed time reversal cavity," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 579 (1992).
[17] M. Tanter, J.-L. Thomas, and M. Fink, "Time reversal and the inverse filter," J. Acoust. Soc. Am. **108**, 223–234 (2000).
[18] M. Tanter, J.-F. Aubry, J. Gerber, J.-L. Thomas, and M. Fink, "Optimal focusing by spatio-temporal inverse filter. I. Basic principles," J. Acoust. Soc. Am. **110**, 37–47 (2001).
[19] J.-F. Aubry, M. Tanter, J. Gerber, J.-L. Thomas, and M. Fink, "Optimal focusing by spatio-temporal inverse filter. II. Experiments," J. Acoust. Soc. Am. **110**, 48–58 (2001).

Revisiting iterative time reversal processing: Application to detection of multiple targets

Gabriel Montaldo,[a] Mickaël Tanter, and Mathias Fink
Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.M.R. C.N.R.S. 7587, 10 rue Vauquelin, 75005 Paris, France

(Received 4 October 2002; accepted for publication 30 October 2003)

The iterative time reversal processing represents a high speed and easy way to self-focus on the strongest scatterer in a multitarget medium. However, finding weaker scatterers is a more difficult task that can be solved by computing the eigenvalue and eigenvector decomposition of the time reversal operator, the so-called DORT method. Nevertheless, as it requires the measurement of the complete interelements response matrix and time-consuming computation, the separation of multiple targets may not be achieved in real time. In this study, a new real time technique is proposed for multitarget selective focusing that does not require the experimental acquisition of the time reversal operator. This technique achieves the operator decomposition using a particular sequence of filtered waves propagation instead of computational power. Due to its simplicity of implementation, this iterative process can be achieved in real time. This high speed selective focusing is experimentally demonstrated by detecting targets through a heterogeneous medium and in a speckle environment. A theoretical analysis compares this technique to the DORT formalism.
 [DOI: 10.1121/1.1636463]

PACS numbers: 43.60.Pt, 43.40.Fz, 43.60.Cg [RLW] Pages: 776–784

I. INTRODUCTION

Over the past decade, acoustic time reversal mirrors[1–3] (TRMs) have been studied widely in medical imaging,[4,5] nondestructive testing,[6] and underwater acoustics.[7,8] An extensive overview of TRM applications can be found in Ref. 9. A time reversal mirror is an array of transducers able to work both in transmitting and receiving mode. When a TRM receives the signal coming from an acoustic source or reflector, it time-reverses the received signals and re-emits them into the medium. The resulting wave front focuses back on the acoustic source or reflector location. When the medium contains several reflectors, the time reversal process can be iterated in order to focus on the brightest target as demonstrated in Refs. 10 and 11 in ultrasonic acoustic laboratory experiments and by Kuperman *et al.* in underwater acoustics.[8,12] The potential of such an iterative time reversal processing was also demonstrated experimentally for medical applications by achieving real time tracking of kidney stones in lithotripsy treatments.[4]

The main limitations of the iterative time reversal process are the impossibility of finding weaker targets and the progressive monochromatization of the signals. At each illumination and reception the signals' bandwidth is narrowed by the limited bandpass of the transducers. Thus, after some iterations, the transmitted and received signals become almost monochromatic.

In order to achieve selective detection and focusing on each scatterer inside an unknown multitarget medium, another approach was developed by Prada *et al.* As is demonstrated in Ref. 13, the echoes of a single target are an eigenvector of the time reversal operator. Using this basic idea that each target is associated to an eigenvector of the time reversal operator it is possible to record the whole time reversal operator and compute its eigenvector decomposition. Thanks to this numerical eigenvector decomposition, a selective focusing on each target can be achieved. Using this technique, known as the DORT method,[13–15] Chambers *et al.* have shown that the spectrum of the time reversal operator can be complex and very informative.[16] However, the DORT method is not a real time method because it requires the measurement of the N×N interelement impulse responses and the computation of the eigenvectors decomposition is quite time consuming.

Recently, Kim *et al.*[17,18] have found an elegant way to detect multiple targets in an underwater waveguide without requiring experimental acquisition of the time reversal operator. The basic idea is first to iterate the time reversal process in order to learn how to focus on the strongest scatterer. Then an algorithm of variance minimization based on the inversion of the cross-spectral density matrix allows one to know how to steer a null on this point. This set of signals avoids any reverberation at the strongest target location and thus can be used to focus on the secondary targets. The second brightest target is then selected by iterating the time reversal process. The prior advantage of this nulling method is that it does not require the fastidious acquisition of the time reversal operator. However, it suffers the same limitation as the iterative time reversal processing: The emission signals' bandwidth is narrowed after each iteration resulting in almost monochromatic signals at the end of the process. Finally, it requires the numerical inversion of the cross-section density matrix. The detection of targets presenting important weight differences is limited by the fact that the strongest

[a]Author to whom all correspondence should be addressed; electronic mail: gabrielmontaldo@yahoo.fr

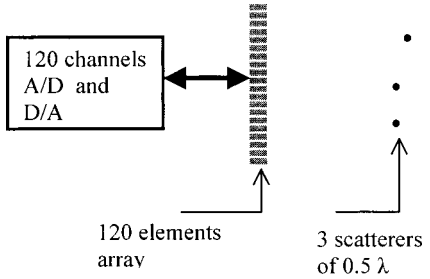


FIG. 1. Configuration of the example. The TRM is composed of an array of 120 elements and a multichannel electronic able to send fully programmable wave forms and record the backscattered echoes. Three targets were embedded in the medium for this example.

target picks up the energy focused on weaker targets and finishes by dominating after some iterations.

In this article, we propose a new real time technique for multitarget selective focusing that does not require the experimental acquisition of the time reversal operator and its numerical decomposition. This technique achieves the operator decomposition simply by using a particular sequence of filtered waves instead of computational power. The general idea of this new approach is to use the time reversal iterative process in order to focus on the brightest target. These signals are then used to derive a broadband cancellation filter which allows one to cancel the echoes of the found target. The selection of the next brightest target is done by iterating the time reversal plus the cancellation filter. This process can be extended to the following of multiple targets detection using the cancellation filter that cleans up the targets already detected. The drawback of temporal spreading induced during the iterative time reversal selection is also overcome by introducing a ''pulse compression operator.'' This operator uses the spread signals coming from the iteration to reconstruct a set of sharp pulsed signals focusing on the selected target.

In Sec. II, the basic principles of the process are explained and illustrated with some preliminary experimental results. In Sec. III, the theoretical formalism of the method is presented and its similarities with the DORT method are highlighted. In Sec. IV, the process is applied to detecting in real time several scatterers through an aberrating medium but also in a noisy ''speckle environment'' as is the case in medical imaging.

II. THE ITERATIVE DECOMPOSITION

In order to explain the iterative decomposition method, we consider a simple homogeneous medium containing three scatterers made with a steel wire of 0.4 mm diameter.

The experiments are conducted in the ultrasonic range in a water tank. The acoustic waves are emitted and received with a standard linear array made of 120 elements working at a 1.5 MHz central frequency. This array is connected to a multichannel system made of 120 independent electronic channels containing A/D and D/A converters (50 MHz sampling, 20 MHz bandwidth). We acquire the signals in a window of 25 μs. The experimental setup is presented in Fig. 1.

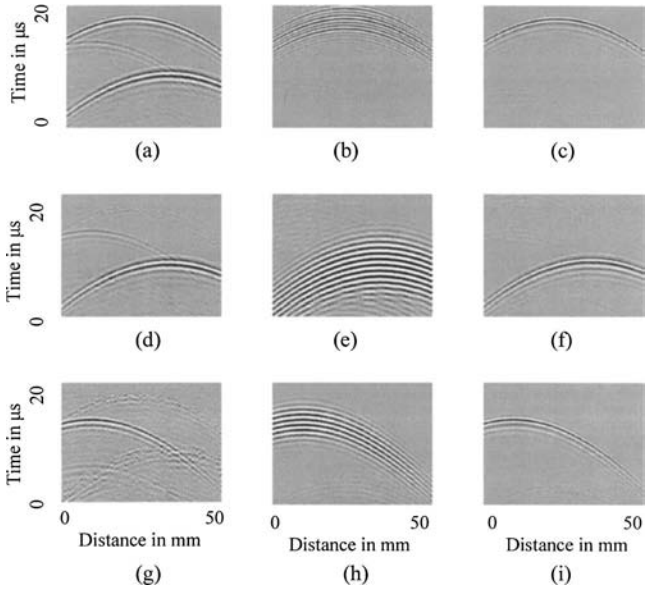


FIG. 2. The iterative process. (a) Backscattered echoes of the three scatterers after a plane wave emission. (b) Detection of the strongest scatterer by iterative time reversal. (c) Eigenpulse of the first scatterer. (d) Backscattered echoes of the two scatterers after filtering the first one. (e) Detection of the second scatterer by time reversal and filtering. (f) Eigenpulse of the second scatterer. (g) Signal of the third scatterer after filtering the first and second one. (h) Detection of the third scatterer by time reversal and filtering. (i) Eigenpulse of the third scatterer.

The process identifies each target step by step beginning from the most reflective one to the weakest one. The selective identification of a new target is made in three steps.

A. First step: Time-reversal processing

A plane wave is emitted in this medium. The backscattered signals are composed of three wave fronts of different amplitudes corresponding to each target [Fig. 2(a)]. If these signals are time reversed and reemitted through the medium, the resulting wave fronts focus on each target and the brightest target is more illuminated than the others. Consequently, its contribution in the backscattered echoes is more important. After a few iterations this time-reversal process permits one to select the most reflecting scatterer. However, at each iteration the signals are filtered by the limited bandwidth of the transducers and it results in a progressive temporal spreading of the emission signals. In Fig. 2(b) we can see the received signal after eight iterations of the time-reversal process, the strongest scatterer was selected but the bandwidth was clearly reduced. The second step of the process allows one to overcome this problem.

B. Second step: Building an "eigenpulse" signal by pulse compression

The bandwidth narrowing suffered during the time reversal process is a real drawback in most applications. An easy solution consists in reconstructing a wideband wave front at each iteration by detecting the arrival time and amplitude law of the signals received on each transducer. The arrival time and amplitude law is then used to reemit a wideband pulsed signal identical on each transducer with the corresponding amplitudes and time delays on each of them. It

FIG. 3. (a) After the iteration of the time reversal process the signal is spread in time. (b) Selection of a wave front. Using a simple algorithm it has a defect. (c) After emitting this wave front, we obtain a narrow pulse without defects in the identification of the wave front.

allows one to avoid the bandwidth spreading of the signals during the iterative process. The arrival time and amplitude can be measured by using a simple maximum detection technique for each transducer (see Fig. 3). In general, the use of a simple algorithm for the ''pulsed wave front'' construction (for example, a maximum detection) can induce some errors in the arrival time estimation. For example, if the signal is composed of several sinusoids as presented in Fig. 3(a) the maximum detection can be limited by a 2π uncertainty [Fig. 2(b)]. It results in the reemission of an incorrect wave front at the next illumination. However, most of the energy of this incorrect wave front is focused on the good location and the maximum detection on the next backscattered echoes becomes easier and more accurate. Thus, a few iterations of the time reversal process combined with the pulse compression allow one to obtain a correct pulsed wave front or ''eigenpulse'' signal [Figs. 3(c) and 2(c)].

In most of our configurations this part of the process was not critical to the convergence of the method. Even if the maximum detection at the first iteration does not give the good wave front [i.e., it has a lot of ''phase gaps'' as in Fig. 3(b)], after the emission of this incorrect wave front a nonnegligible part of the energy focuses on the target and the backscattered echoes are short pulses for which the maximum detection gives a better wave front.

C. Third step: Canceling the detected scatterer and selecting a new one

The basic idea for selecting a new scatterer is to apply a ''cancellation operator'' able to clear up all the signals coming from the previous detected target. The details of this operator will be explained in Sec. III. If we start with a backscattered signal containing the echoes of the three scatterers [Fig. 2(a)], after applying the cancellation operator we obtain the filtered signals shown in Fig. 2(d). As we can see, the echoes of the strongest scatterer have been canceled. This new set of filtered signals is now used as initial illumination for the iterative time reversal process. The cancellation filter is applied at each step during the iterative time reversal process proposed in Secs. II A and II B. Consequently the first target is not illuminated, the second target generates the brightest echoes and it is progressively selected by the iteration process [Fig. 2(e)].

The signals backscattered by the second target are temporally spread and can be ''pulse compressed'' as described in Sec. II B [Fig. 2(f)]. Finally, the cancellation filter permits one to cancel the first and second targets and selects the third target by iterating the time reversal process. Figures 2(g)–(i) describe the final eigenvector decomposition that was found for the third and weakest target.

The complete process does not require any hard numerical computation, the cancellation and pulse compression operators are easy to calculate, and the combination of steps II B and II C was found sufficient to select multiple targets. The speed of detection is directly linked to the traveling time of the waves in the medium. For short distances as in the human body, this method can be implemented for real time selective focusing. For example with a propagation distance of 15 cm a step of iterative time reversal process takes only 1 ms, the maximum detection and the cancellation filter can be done in 3 ms. Running the complete detection of the three targets of Fig. 1 takes nearly 50 ms. With this short time of detection it is possible to track moving multiple targets.

We will demonstrate experimentally in Sec. IV that this process is robust not only in a noisy environment (like speckle noise for medical applications) but also through an aberrating medium. For this purpose, this technique can also be seen as a phase aberration correction technique using the bright spots embedded in the medium.

III. THEORETICAL APPROACH

A. The iteration of the time reversal

We consider a system composed of an array of transducers capable of working both in transmitting and receiving mode and a linear medium of propagation containing a few point-like scatterers. The propagation medium is heterogeneous but we suppose that the amplitude and phase perturbation are not very different for different frequencies (i.e., strongly dispersive media will not be considered).

The signals received and emitted by the transducers array are functions of space and time $e=e(x,t)$, where x denotes the different positions of the individual transducers.

A complete time reversal iteration has two steps: first, a signal $e=e(x,t)$ is emitted and the echoes backscattered by the medium $r=r(x,t)$ are recorded by the array. Second, these signals are time reversed $r(x,-t)$ and they are reemitted by the array. The backscattered echoes are recorded by the array and correspond to the first set of iterated signals $e^{(1)}=e^{(1)}(x,t)$. The original set of signals and its first iteration are related by the time reversal operator of the medium $\mathbf{H}$ and can be written in a compact form

$$e^{(1)}=\mathbf{H}e. \tag{1}$$

Note that $\mathbf{H}$ denotes the application of a spatiotemporal linear operator to the signals $e(x,t)$. The notation H will refer to the matrix representation of this operator. The eigenmodes of this time reversal operator $\mathbf{H}$ are a set of signals $\nu_i=\nu_i(x,t)$ that verify $\mathbf{H}\nu_i=\lambda_i\nu_i$, where λ_i is the associated eigenvalue. In a free space medium with some punctual targets, the eigenvectors are monochromatic cylindrical waves focused on the targets. Assuming that the sound speed and absorption heterogeneities are due to a random screen located closed to the array, it introduces only different amplitude gains and time delays on each transducer of the array. Then the eigenmodes can be written as

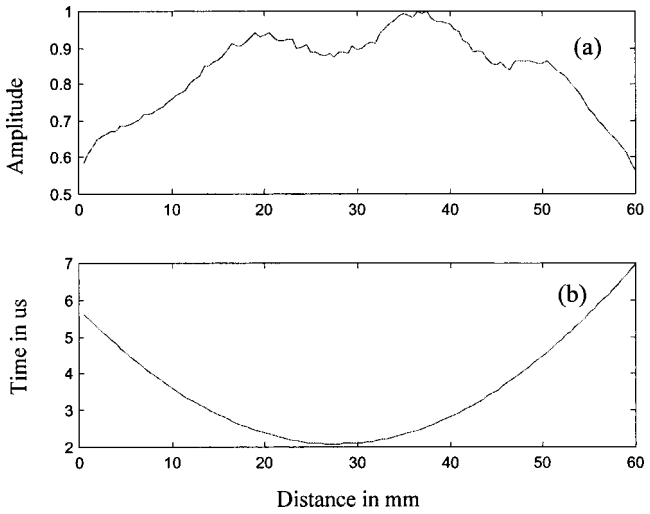


FIG. 4. (a) Amplitude $A(x)$ and delay $\tau(x)$ for the first scatterer of Fig. 1.

$$\nu_{i\omega}(x,t)=A_i(x)\exp(j\omega[t+\tau_i(x)]), \tag{2}$$

where $\tau_i(x)$ is a time delay law depending only on the scatterer location and medium aberrations and $A_i(x)$ is the amplitude of the signal coming from the scatterer and received on the transducer located at x. $A_i(x)$ and $\tau_i(x)$ describe completely the eigenmode associated with scatterer i. As an example, Fig. 4 represents the experimental amplitude and time delay law associated with the first target corresponding to the Fig. 1 setup. An eigenmode associated with this particular target at 1 MHz frequency is presented in Fig. 5.

Note that the eigenmodes $\nu_{i\omega}(x,t)$ and their eigenvalues $\lambda_{i\omega}$ must validate

$$\mathbf{H}\nu_{i\omega}=\lambda_{i\omega}\nu_{i\omega}, \tag{3}$$

and the orthogonal relationship

$$\int\int \nu^*_{i\omega}(x,t)\nu_{j\omega'}(x,t)dx\,dt=\delta_{ij}\delta_{\omega\omega'}. \tag{4}$$

We can build a matrix representation of the time reversal operator in the base of the spatiotemporal eigenmodes as

$$H_{i\omega,j\omega'}=\int\int \nu^*_{i\omega}(x,t)\mathbf{H}\nu_{j\omega'}(x,t)dx\,dt. \tag{5}$$

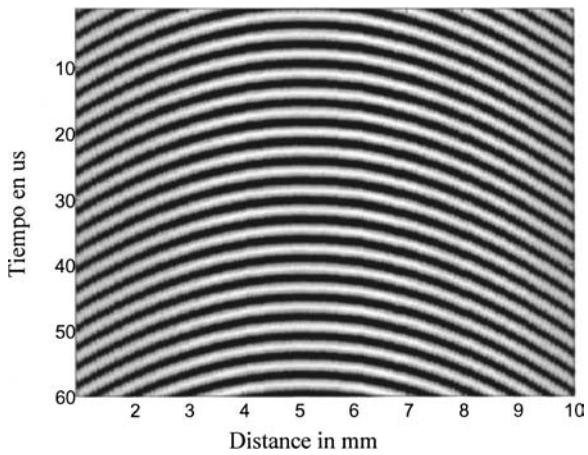


FIG. 5. Eigenmode of the first scatterer for a frequency of 1 MHz.

This matrix is diagonal and contains the eigenvalues $\lambda_{i\omega}$. If we rank the pairs of index i, ω first by the scatterer index i and after by its frequency ω, we can see this matrix in a graphical representation as

$$H_{i\omega,j\omega'}=\begin{bmatrix} \lambda_{11} & \cdot & \cdot & 0 \\ & \lambda_{1\ldots} & & \\ & & \lambda_{1n} & \\ & & & \lambda_{21} \\ & & \lambda_{2\ldots} & \\ 0 & & & \lambda_{2n} \end{bmatrix} \tag{6}$$

In this case we assumed that there are only two scatterers in the medium. Using this representation of the time reversal operator, it is easy to see what happens in an iteration of the time reversal process. We start with an arbitrary excitation $e(x,t)$ and decompose it in the eigenvector base

$$e(x,t)=\sum_{i\omega} e_{i\omega}\nu_{i\omega}(x,t), \tag{7}$$

where $e_{i\omega}$ are the weights of each eigenvector. After applying the time reversal operator, the new coefficients of the signals are $e^{(1)}_{i\omega}=\lambda_{i\omega}e_{i\omega}$ and after n iterations the coefficients become $e^{(n)}_{i\omega}=\lambda^n_{i\omega}e_{i\omega}$. After each iteration, the coefficient of the strongest eigenvalue is strengthened in comparison to the secondary scatterers and the time reversal converges toward the eigenmode corresponding to the strongest eigenvalue.

As the response of the transducer and the cross section of the scatterers depend on the frequency, $\lambda_{i\omega}$ has different values for different frequencies. After each iteration, the coefficient of the strongest eigenvalue is strengthened and the time reversal converges toward the eigenmode corresponding to the strongest eigenvalue. Experimentally, the bandwidth of the signal is narrowed at each step of the iterative process and the signal spreads in time.[19]

B. Iteration with an intermediate operator

As was previously seen, the iterative time reversal permits one to identify the brightest scatterer embedded in a medium but it presents some inconveniences. First, it is not possible to find weaker targets. Second, as the eigenvectors are monochromatic waves, the iterative time reversal expands the length of the signals and it is not possible to apply a large number of iterations. A modification of the time reversal process is introduced to overcome these limitations. The main idea is to apply an intermediate operator $\mathbf{O}$ after each iteration of the time reversal process. Thus, if we start with a set of emission signals $e(x,t)$, the new set of signals after the iteration will be

$$e^{(1)}=\mathbf{HO}e=\tilde{\mathbf{H}}e, \tag{8}$$

where $\tilde{\mathbf{H}}=\mathbf{HO}$ is the effective iteration operator. We can define the matrix representation of $\mathbf{O}$ in the base of eigenvectors of $\mathbf{H}$ in a similar manner to Eq. (5),

$$O_{i\omega,j\omega'}=\int\int \nu^*_{i\omega}(x,t)\mathbf{O}\nu_{j\omega'}(x,t)dx\,dt, \tag{9}$$

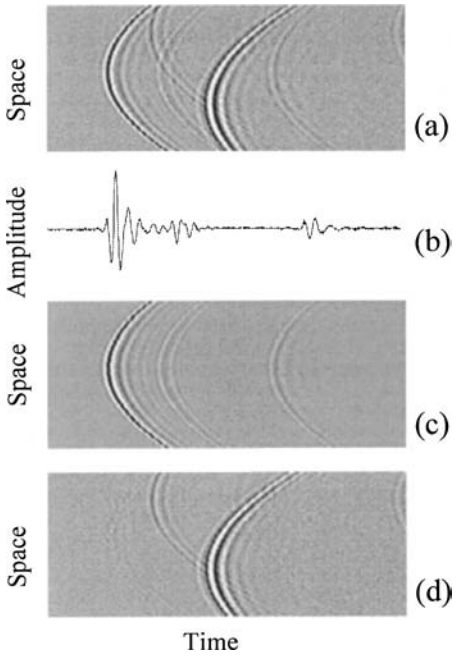


FIG. 6. Illustration of the filter operator. From the original image (a) we calculate the weight of the first scatterer at each instant with Eq. 13(b). Performing the convolution of this weight with the wave front function we obtain the echoes of the first scatterer (c). Subtracting (c) from (a) we obtain the echoes of the other scatters (d).

Decomposing the signal $e(x,t)$ as in Eq. (7) it is easy to calculate that the coefficients $e_{i\omega}^{(1)}$ after the iteration are

$$e_{i\omega}^{(1)} = \sum_{j\omega'} \tilde{H}_{i\omega,j\omega'} e_{j\omega'}, \tag{10}$$

where

$$\tilde{H}_{i\omega,j\omega'} = \sum_{k\omega''} H_{i\omega,k\omega''} O_{k\omega'',j\omega'}. \tag{11}$$

An eigenvector $b_{i\omega}$ of the $\tilde{H}_{i\omega,j\omega'}$ matrix associated with an eigenvalue β verifies $\Sigma_{j\omega'} \tilde{H}_{i\omega,j\omega'} b_{j\omega'} = \beta b_{i\omega}$. After iterating n times, the process amplifies β^n times this eigenvector. The strongest eigenvalue is finally selected after some iterations of Eq. (10). The use of an adequate operator **O** allows us to transform the convergence of the time reversal process from the strongest eigenvalue of **H** to the strongest eigenvalue of $\tilde{\mathbf{H}}$.

C. The cancellation operator

This operator solves the first problem of finding weaker scatterers in a multitarget environment, it permits one to clean the echoes of a brighter scatterer previously detected in order to find the next one.

The calculus of this operator starts with the knowledge of the time delay $\tau_i(x)$ and the amplitude $A_i(x)$ of the scatterer, with these data we can build a ''wave form function'' as

$$w_1(x,t) = \delta[t + \tau_1(x)] A_1(x). \tag{12}$$

where δ is the Dirac's function. In this example we chose the wave form function for the first scatterer.

If we have a signal $e(x,t)$, we can build a function $P(t)$ that gives the weight of the echoes of the first scatterer in the signal at each time as

$$P(t) = \int \int e(x',t') w_1(x',t'-t) dt'\, dx'$$
$$= \int e[x', t - \tau_1(x')] A_1(x') dx'. \tag{13}$$

At each time t, the multiplication by the wave form function selects the echoes of the first scatterer at time t. Performing the integration we obtain an amplitude of the echoes of the first scatterer [Fig. 6(b)].

With the amplitude we can rebuild the echoes of the first scatterer performing the correlation between $P(t)$ and the wave form function [see Fig. 6(c)],

$$\int P(t') w_1(x, t-t') dt' = A_1(x) P[t + \tau_1(x)]. \tag{14}$$

Substituting $P(t)$ by Eq. (13), the echoes of the first scatterer are

$$A_1(x) \int A_1(x') e[x', t + \tau_1(x) - \tau_1(x')] dx'.$$

Finally we subtract these echoes from the signal $e(x,t)$ to obtain one without echoes from the first scatterer [Fig. 6(d)]. As a conclusion the cancelation operator is

$$\mathbf{F}^1 e(x,t) = e(x,t) - A_1(x) \int A_1(x') e[x', t + \tau_1(x) - \tau_1(x')] dx'. \tag{15}$$

Calculating the elements of this operator with the matrix formalism introduced in Eq. (9), we obtain

$$F^1_{i\omega,j\omega'} = (1 - \delta_{1i}) \delta_{ij} \delta_{\omega\omega'}. \tag{16}$$

The complete deduction of Eq. (16) is given in the appendix. Note that the diagonal elements of the matrix defined in Eq. (16) are equal to 1 except for all the frequency components of the first target which are equal to zero. After multiplying by the diagonal matrix $H_{i\omega,j\omega'}$, we obtain the matrix representation of the operator $\mathbf{HF}^1$,

$$HF^1=\begin{bmatrix}\lambda_{11} & \cdot & \cdot & 0\\ & \lambda_{1\ldots} & & \\ & & \lambda_{1n} & \\ & & & \lambda_{21}\\ & & & & \lambda_{2\ldots}\\ 0 & & & & & \lambda_{2n}\end{bmatrix}\times\begin{bmatrix}0 & & & & & 0\\ & 0 & & & & \\ & & 0 & & & \\ & & & 1 & & \\ & & & & 1 & \\ 0 & & & & & 1\end{bmatrix}=\begin{bmatrix}0 & & & & & 0\\ & 0 & & & & \\ & & 0 & & & \\ & & & \lambda_{21} & & \\ & & & & \lambda_{2\ldots} & \\ 0 & & & & & \lambda_{2n}\end{bmatrix}.$$

Due to the cancellation of the diagonal values of the first eigenvector, the second scatterer presents the strongest eigenvalue in HF^1, and iterating the operator $\mathbf{HF}^1$, it will converge to the second brightest scatterer.

D. The pulse compression operator

One problem of the iteration with the time reversal is that the signals are enlarged in each iteration. In Sec. II B we proposed to send only one wave front in order to compress the signals. The wave front can be described by the function (12). We can estimate the time delay $\tau_i(x)$ and the amplitude $A_i(x)$ using a maximum detection at each point x. With these estimates it is possible to build a contracting filter that generates a wave front multiplied by the weight depending on the signal $e(x,t)$ as

$$\mathbf{M}^1 e(x,t)=w_1(x,t)\int\int w_1(x',t')e(x',t')dx'\,dt'. \tag{17}$$

The integral $\int\int w_1(x',t')e(x',t')dx'dt'$ gives the weight of the wave front in the signal $e(x,t)$ at time $t=0$ as in Eq. (13).

After performing the integration and substituting $w_1(x,t)$ by Eq. (12) we arrived at a simple expression for the contracting operator,

$$\mathbf{M}^1 e(x,t)=A_1(x)\,\delta[t+\tau_1(x)]\times\int A_1(x')e[x',-\tau_1(x')]dx'. \tag{18}$$

Calculating the projected matrix for this operator as in Eq. (9) we obtain

$$M^1_{i\omega j\omega'}=\delta_{1i}\delta_{1j}1_{\omega\omega'}, \tag{19}$$

where $1_{\omega,\omega'}$ is a matrix of 1 for all values of ω and ω'. The complete calculus of Eq. (19) is given in the appendix. This operator can be seen as a mask set to 1 in a square corresponding to all frequency components of the first scatterer and 0 elsewhere. After the multiplication of the mask matrix by the diagonal matrix $H_{i\omega,j\omega'}$, we obtain the matrix representation of the operator $\mathbf{HM}^1$,

$$HM^1=\begin{bmatrix}\lambda_{11} & \cdot & \cdot & 0\\ & \lambda_{12} & & \\ & & \lambda_{13} & \\ & & & \lambda_{21}\\ & & & & \lambda_{22}\\ 0 & & & & & \lambda_{23}\end{bmatrix}\times\begin{bmatrix}1&1&1&0&\cdot&0\\1&1&1&0&\cdot&\cdot\\1&1&1&0&\cdot&\cdot\\0&0&0&0&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&0&\cdot\\0&\cdot&0&\cdot&\cdot&0\end{bmatrix}=\begin{bmatrix}\lambda_{11}&\lambda_{11}&\lambda_{11}&0&\cdot&0\\\lambda_{12}&\lambda_{12}&\lambda_{12}&0&\cdot&\cdot\\\lambda_{13}&\lambda_{13}&\lambda_{13}&0&\cdot&\cdot\\0&0&0&0&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&0&\cdot\\0&\cdot&0&\cdot&\cdot&0\end{bmatrix}.$$

The eigenvector of highest eigenvalue is $(\lambda_{11},\lambda_{1,\ldots},\lambda_{1n},0,\ldots,0)$ and its eigenvalue is

$$\tilde{\lambda}_i=\sum_{\omega}\lambda_{i\omega}. \tag{20}$$

In conclusion, the compression operator allows one to select all the frequency components of a same scatterer and add them. As all the frequency components are summed the

TABLE I. Scheme of the applied operators and their function.

Step	Operator	Action
1	H^n	After n iterations of time reversal processing, the first target is selected. τ_1 and A_1 are estimated.
2	HM^1	Building the first "eigenpulse." τ_1 and A_1 are then reevaluated.
3	$(HF^1)^n$	Filtering the first target, the second becomes the strongest. τ_2 and A_2 are estimated after n iterations.
4	$(HF^1)^nM^2$	Building the second eigenpulse. τ_2 and A_2 are reevaluated.
5	$(HF^2F^1)^n$	Filtering the first and second scatter. The third becomes the strongest. τ_3 and A_3 are estimated after n iterations.
6	$(HF^2F^1)^nM^3$	Building the third eigenpulse. τ_3 and A_3 are reevaluated.
7	$(HF^3F^2F^1)^n$	Filtering the first, second, and third targets. The fourth becomes the strongest ...

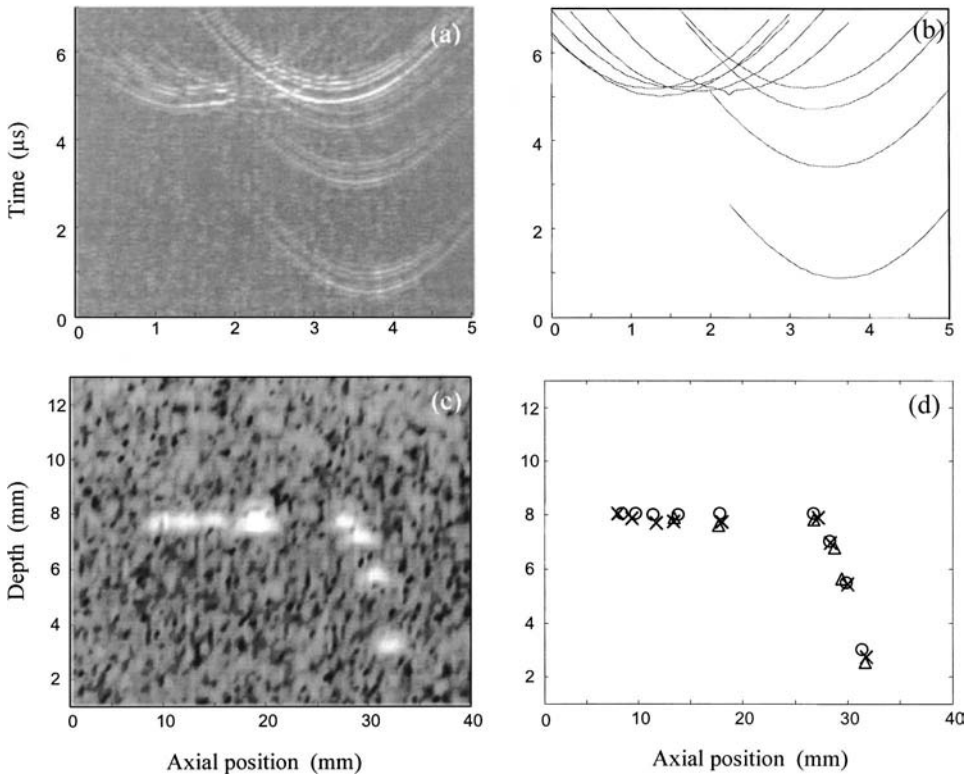


FIG. 7. Target detection in a biological phantom. (a) Signal received after a plane wave illumination, we can see some echoes of the target with an important noise of the structure of the phantom. (b) Detection of the echoes of the nine targets. (c) Echographic image of the phantom due to the speckle noise. The first three targets are difficult to resolve. (d) Calculation of the target position from the detected echoes. Crosses are the measured positions, the circles are given by the furnisher of the phantom, and the triangles are the positions calculated from the maximum of the *B* scan. The maximum detection in the *B* scan is not able to give the positions and number of scatters for the first three scatterers well.

temporal spreading is avoided and a narrow pulse that focuses on the target is possible to obtain.

This technique is similar to applying an inverse filter to compensate the bandwidth of the transducer. However, such an inverse filter is more sensitive to noise, needs a precalibration of the transducers and lots of calculations that limit the real-time aspect of the technique. Extending these two operators $\mathbf{F}^1$ and $\mathbf{M}^1$ to other targets we obtain two sets of operators: the filter operators $\mathbf{F}^i$ which cancel the echoes of the ith scatterer and the pulse compression operators $\mathbf{M}^i$ which compact the echoes of scatterer i in a narrow pulsed signal. These two kinds of operators are used to detect all the targets successively as summarized in Table I. As a final observation, this system has some similitude with the DORT method.

The DORT method starts with the measurement of the interelement response matrix. This matrix is a representation of the time reversal operator **H**. Performing a numerical decomposition of this matrix we obtain the complete set of eigenmodes $\nu_{i\omega}(x,t)$. Adding all the eigenmodes ω of a same target i we can build a pulse that focuses on the target as $\nu_i(x,t)=\Sigma_\omega \nu_{i\omega}(x,t)$.

However, in the DORT method the decomposition is achieved frequency by frequency and an eigenpulse can be built by recombining the DORT eigenvectors but there is a relative phase ambiguity between eigenvectors of different frequencies. This problem makes the recombination of the eigenvectors nonunique.

This final result is similar to the ''eigenpulses'' obtained by the iterative decomposition but substituting the numerical calculation by a specific convergence of the time reversal process over each target.

IV. EXAMPLES OF DETECTION IN INHOMOGENEOUS MEDIA

The most interesting applications of multitarget detection correspond to inhomogeneous media. If the phase and amplitude aberration are not extremely hard we are in the hypotheses that $\tau_i(x)$ and $A_i(x)$ are independent of the frequency and this method is able to detect the targets.

The first application is to find nine nylon wires of half wavelength diameter (0.4 mm) in a biological phantom with an important ''speckle noise'' structure.

Figure 7(a) shows the echoes of the phantom when it is illuminated with a plane wave. We can see the echoes of some targets superposed to a structural speckle noise coming from a random distribution of small scatterers. The iterative method is able to identify the nine echoes as is shown in Fig. 7(b).

Due to the speckle noise a standard echographical image is not able to resolve the first three scatterers of the left-hand side of the image [Fig. 7(c)]. Using the time delays detected

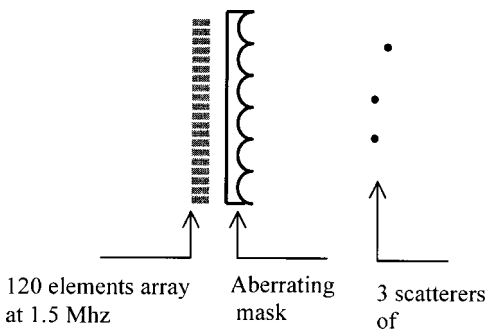


FIG. 8. Identification through an aberrating mask.

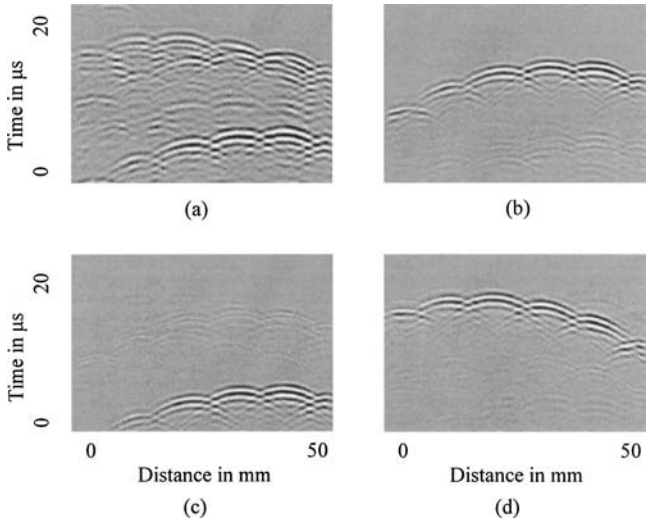


FIG. 9. Identification of three scatterers through an aberrating mask. (a) Echoes of the three scatters after a plane wave excitation. (b), (c), and (d) Identification of the three scatters.

by the iterative method we can calculate the position of each scatterer. Figure 7(d) shows the calculated positions, the positions given by the furnisher and the ones obtained taking the maximum of the *B* scan of Fig. 7(c).

The maximum detection in the *B*-scan image can be applied to the last six scatterers where the resolution is enough to isolate each target. However, for the first three scatterers it is impossible to give the good number of targets and their respective positions directly from the classical *B*-scan image.

A possible application could be the identification of microcalcifications in the breast, if a suspicious area is detected by a conventional image, an automatic detection of punctual scatterers could give extra information to detect the calcifications.

The final experience consisted in putting an aberrating mask between the array and the scatterers. The mask is made of a soft polymer composed of semicircles of 1 cm each, the scatterers are the same as in the experiment of Fig. 1 (see Fig. 8). Figure 9(a) shows the echoes of three scatterers when we insonify them with a plane wave.

In Fig. 10 we can see the *B* scan of the three scatterers through the aberrating mask. Due to the particular geometry

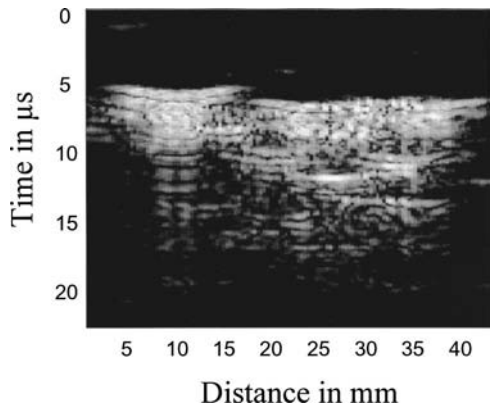


FIG. 10. *B* scan of the scatterers through the mask, there are a lot of ''virtual images'' of the scatters that do not permit to identify the scatters.

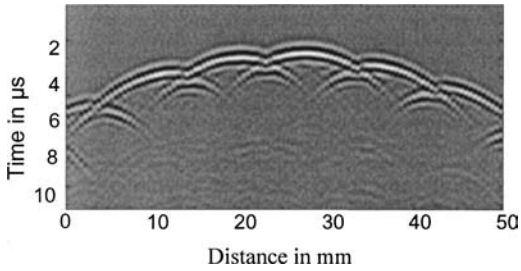


FIG. 11. Simulation of the echo of a single target through the aberrating mask used in the experience of Fig. 9. We can see the resemblance between this simulation and the eigenpulses detected in Fig. 9.

of the mask the *B* scan shows a lot of ''virtual replica'' of the scatterers, and then it is impossible to know how many real scatterers there are in the medium and how to focus on each one.

After the application of the iterative method the three scatterers were identified [Figs. 9(b)–(d)]. The wave fronts are composed of a succession of small arches. We have done a simulation of the aberrator by finite differences in order to verify the form of the eigenmodes. In Fig. 11 we show the simulation of the first scatterer that has the same form as Fig. 9. This last example clearly demonstrates the efficiency of the method in a strongly aberrating medium. These experimental examples show a good stability of the method. A failure of this iterative method was detected when two targets of very different size were located close together. In this case, the cancellation operator is not precise enough to clean up completely the echoes of the strongest target and consequently the detection oscillates between the two targets.

V. CONCLUSION

A new real time technique has been proposed for selective focusing on multiple targets. It is based on the combination of the iterative time reversal process with very simple operations like maximum detection and signal subtraction. This method enables one to find and auto focus in real time on multiple targets using two intermediate operators. The cancellation operator permits one to cancel the echoes coming from targets already detected and the iterative time reversal process achieves the adaptive focusing on a new target. A pulse compression operator based on the use of a maximum detection algorithm serves to overcome the problem of signal temporal spreading induced during the iterative time reversal process. Theoretical similarities with the DORT method have been highlighted. Instead of achieving the eigenvector decomposition using time consuming computational power, the new approach builds it experimentally in real time by simply using wave propagation. This high speed selective focusing technique has a potential application in microcalcification detection in breast, nondestructive testing and underwater acoustics. For all these topics the real time capability could be interesting for tracking moving multiple targets.

APPENDIX

We consider a simplified medium containing only a discrete number of scatterers. The aberrations induced by the medium are assumed to be restricted to an amplitude and time delay law independent of the frequency. This amplitude and time delay law differ from one scatterer to the other. The eigenmodes associated with a target are monochromatic waves with an amplitude and phase perturbation assumed to be constant over the transducer bandwidth,

$$\nu_{i\omega}(x,t)=A_i(x)\exp(j\omega[t+\tau_i(x)]), \tag{A1}$$

where the index i denotes the target number and ω is the frequency of the eigenmode. $\tau_i(x)$ is a time delay depending on the array location of the elements of the array. $A_i(x)$ is the amplitude of the signal received on each spatial point. Thus, $\tau_i(x)$ and $A_i(x)$ fully characterize the wave front coming from the target i. The eigenmodes must validate the orthogonal relationship

$$\int\int \nu^*_{i\omega}(x,t)\,\nu_{j\omega'}(x,t)dx\,dt=\delta_{ij}\delta_{\omega\omega'}. \tag{A2}$$

This calculation introduces a condition between the amplitudes and the time delay

$$\int A_i(x)A_j(x)\exp(j\omega[\tau_j(x)-\tau_i(x)])dx=\delta_{ij}. \tag{A3}$$

With this relation we can calculate the matrix representation of the operators.

1. The cancellation operator

The action of the cancellation operator over a signal $e(x,t)$ was defined in Eq. (15) as

$$\mathbf{F}^1e(x,t)=e(x,t)-A_1(x)\int e[x',t+\tau_1(x)-\tau_1(x')]\times A_1(x')dx'. \tag{A4}$$

When we apply this filter to an eigenmode $\nu_{i\omega}$ and use Eq. (A3) the integral becomes

$$A_1(x)\int \nu_{i\omega}[x',t+\tau_1(x)-\tau_1(x')]A_1(x')dx' = \nu_{i\omega}(x,t)\,\delta_{1i}. \tag{A5}$$

The action of the filter over an eigenvector is

$$\mathbf{F}^1\nu_{i\omega}=(1-\delta_{1i})\,\nu_{i\omega}. \tag{A6}$$

The matrix representation of the $\mathbf{F}$ operator is calculated as

$$F^1_{i\omega,j\omega'}=\int\int \nu^*_{i\omega}(x,t)\mathbf{F}^1\nu_{j\omega}(x,t)dx\,dt. \tag{A7}$$

Using Eq. (A7) and the orthogonal relationship (A2) we obtain

$$F^1_{i\omega,j\omega'}=(1-\delta_{1i})\,\delta_{ij}\delta_{\omega\omega'}. \tag{A8}$$

2. The eigenpulse mask operator

We have defined the action of the mask operator over a signal $e(x,t)$ as

$$\mathbf{M}^1e(x,t)=A_1(x)\,\delta[t+\tau_1(x)]\int A_1(x')\times e[x',-\tau_1(x')]dx'. \tag{A9}$$

Using Eq. (A3), the action over an eigenvector $\nu_{j\omega'}$ is

$$\mathbf{M}^1\nu_{j\omega'}(x,t)=A_1(x)\,\delta[t+\tau_1(x)]\,\delta_{1j}. \tag{A10}$$

The matrix representation is calculated as

$$M^1_{i\omega,j\omega'}=\int\int \nu^*_{i\omega}(x,t)\mathbf{M}^1\nu_{j\omega'}(x,t)dx\,dt, \tag{A11}$$

using the orthogonal relationship (A2) we obtain

$$M^1_{i\omega,j\omega'}=\delta_{1i}\delta_{1j}. \tag{A12}$$

This operator is independent of the frequencies ω and ω'. In order to show this propriety we multiplied Eq. (16) by $1_{\omega\omega'}$.

[1] O. Ikeda, "An image reconstruction algorithm using phase conjugation for diffraction-limited imaging in an inhomogeneous medium," J. Acoust. Soc. Am. **85**, 1602–1606 (1989).
[2] M. Fink, "Time-reversal mirrors," J. Phys. D **26**, 1330–1350 (1993).
[3] M. Fink, "Time reversal of ultrasonic fields. II. Basic Principles," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 554–566 (1992).
[4] J. L. Thomas, F. Wu, and M. Fink, "Time reversal mirror applied to litotripsy," Ultrason. Imaging **18**, 106–121 (1996).
[5] M. Tanter, J.-L. Thomas, and M. Fink, "Focusing and steering through absorbing and aberating layers: Application to ultrasonic propagation through the skull," J. Acoust. Soc. Am. **103**, 2403–2410 (1998).
[6] N. Chakroun, M. Fink, and F. Wu, "Time reversal processing in ultrasonic non destructive testing," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**, 1087–1098 (1995).
[7] D. R. Jackson and D. R. Dowling, "Phase conjugation in underwater acoustics," J. Acoust. Soc. Am. **89**, 171–181 (1991).
[8] W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, "Phase conjugation in the ocean: Experimental demonstration of an acoustic time reversal mirror," J. Acoust. Soc. Am. **103**, 25–40 (1998).
[9] M. Fink, D. Cassereau, A. Derode, C. Prada, P. Roux, M. Tanter, J. L. Thomas, and F. Wu, "Time-reversed acoustics," Rep. Prog. Phys. **63**, 1933–1995 (2000).
[10] C. Prada, F. Wu, and M. Fink, "The iterative time reversal process: A solution to self-focusing in the pulse echo mode," J. Acoust. Soc. Am. **90**, 1119–1129 (1991).
[11] C. Prada, J. L. Thomas, and M. Fink, "The iterative time reversal process: Analysis of the convergence," J. Acoust. Soc. Am. **97**, 62–71 (1995).
[12] H. C. Song, W. A. Kuperman, and W. S. Hodgkiss, "Iterative time reversal in the ocean," J. Acoust. Soc. Am. **105**, 3176–3184 (1999).
[13] C. Prada and M. Fink, "Eigenmodes of the time reversal operator: A solution to selective focusing in multiple-target media," Wave Motion **20**, 151–163 (1994).
[14] C. Prada, S. Manneville, D. Spoliansky, and M. Fink, "Decomposition of the time reversal operator: Detection and selective focusing on two scatterers," J. Acoust. Soc. Am. **99**, 2067–2076 (1996).
[15] N. Mordant, C. Prada, and M. Fink, "Highly resolved detection and selective focusing in a waveguide using the D.O.R.T. method," J. Acoust. Soc. Am. **105**, 2634–2642 (1999).
[16] D. H. Chambers and A. K. Gautesen, "Time reversal for a single spherical scatterer," J. Acoust. Soc. Am. **109**, 2616–2624 (2001).
[17] J. S. Kim, H. C. Song, and W. A. Kuperman, "Adaptive time reversal mirror," J. Acoust. Soc. Am. **109**, 1817–1825 (2001).
[18] J. S. Kim, W. S. Hodgkiss, W. A. Kuperman, and H. C. Song, "Null-broadening in a waveguide," J. Acoust. Soc. Am. **112**, 189–197 (2002).
[19] M. Cheney, D. Isaacson, and M. Lassas, "Optimal acoustics measurements," SIAM (Soc. Ind. Appl. Math.) J. Appl. Math. **61**, 1628–1647 (2001).

Data extrapolation method for boundary element method-based near-field acoustical holography

Kenji Saijyou[a] and Hiroshi Uchida[b]
Fifth Research Center, Technical R & D Institute, Japan Defense Agency, 3-13-1 Nagase, Yokosuka, 239-0826, Japan

(Received 8 May 2003; revised 3 November 2003; accepted 17 November 2003)

A new data extrapolation method for boundary element method (BEM)-based near-field acoustical holography (NAH) is proposed to reduce an error of the reconstructed result obtained from the pressure measured on an aperture small compared with the structure. The finiteness of the measurement aperture is a serious impediment to actual large-scale implementation of NAH, because NAH requires the measurement of the pressure field over a complete surface of structure. To eliminate the requirement, the wave number-space data extrapolation method for fast Fourier transform (FFT)-based NAH has been proposed. In this paper, the extension of this data extrapolation method to BEM-based NAH is presented. The effectiveness of this method is demonstrated by experiments. The experiment results confirm that the reconstruction error is sufficiently suppressed by the proposed method. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1641377]

PACS numbers: 43.60.Pt, 43.60.Sx, 43.20.Tb, 43.40.Rj [EJS] Pages: 785–796

I. INTRODUCTION

As a method for analysis of structural acoustics, near-field acoustical holography (NAH) is widely used nowadays. In this technique, the normal velocity on a surface of a structure is reconstructed from the pressure measured on a near-concentric hologram surface. A two-dimensional fast Fourier transform (FFT) has been used for simple shape structures such as plates, cylinders, and spheres.[1–4] The algorithm of FFT-based NAH is simple and efficient. However, FFT-based NAH has been limited to structural surfaces with simple geometries, for example, planar and cylindrical surfaces. This requirement has not always been met in many industrial applications where complex geometrical structures have been present.

In order to eliminate this requirement, boundary element method (BEM)-based NAH has been studied by many researchers.[5–14] In BEM-based NAH, geometric and vibro-acoustic relation between the vibrating structure and the measurement hologram is modeled as the vibro-acoustic transfer matrix by means of BEM. And, the distribution of the surface velocity of the structure can be reconstructed by multiplying the inverse of the vibro-acoustic transfer matrix. For practical implementation of BEM-based NAH, appropriate regularization is necessary to reduce the reconstruction error, because this problem is an ill-posed inverse problem due to the existence of strongly decaying evanescent-like waves. Therefore, the transfer matrix is decomposed by singular value decomposition (SVD), and an appropriate regularization filter is applied to suppress the amplification of measurement noise.

Although NAH is a very powerful tool for structural acoustics, there is an obstacle which makes it difficult to implement NAH to actual large-scale structure. NAH requires the measurement of the pressure field over a complete surface located exterior to the structure. Therefore, the finiteness of the measurement aperture is a serious impediment for applying NAH to the localization of the noise source to field measurement. In FFT-based NAH, wraparound error is induced by the periodic characteristic of FFT. To reduce the wraparound error, zero padding is applied to the measured pressure hologram for virtual enlargement of the hologram.[1,4] If the measurement aperture is sufficiently large, the pressure at the edge of the measurement aperture is negligibly small, and zero padding reduces the reconstruction error. However, if the pressure at the edge of the measurement aperture has to be considered, zero padding creates a discontinuity of the pressure. This discontinuity causes a tremendously large error. BEM-based NAH is free from the wraparound error. However, the discontinuity of the aperture edge contaminates the reconstructed result similarly.

To suppress the influence of the finiteness of the measurement aperture, the K-space data extrapolation method[15] has been proposed, which is applicable for FFT-based NAH. In this paper, this data extrapolation method is extended and applied to BEM-based NAH. The noise contained in the measured pressure contaminates the reconstructed result, especially in the field measurement. To suppress the influence of noise, an appropriate regularization is necessary. Some regularization methods have been proposed. However, these methods cannot directly estimate an appropriate regularization filter, because the estimation of the appropriate regularization filter is based on the wave number-space (abbreviated as ''K space,'' hereafter) spectrum of the pressure, and the finiteness of the measurement aperture deforms the spectrum. Therefore, an estimation method of an appropriate regularization filter is also presented.

Recently, a similar approach has been published.[16] In this method, the determination method of the regularization

[a]Electronic mail: saijyou@jda-trdi.go.jp
[b]Electronic mail: hutida@jda-trdi.go.jp

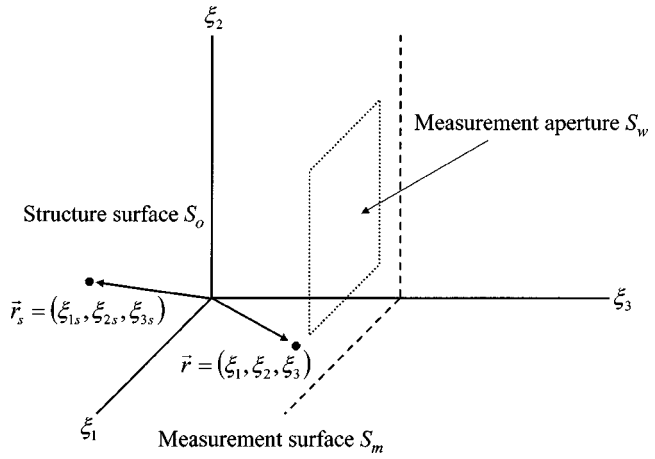


FIG. 1. The general coordinate system applied to the data extrapolation method.

parameter is based on Morozov discrepancy principle (MDV). The MDV provides a simple method to find the regularization parameter when the variance of the noise is known. However, it is difficult to obtain the noise variance from zero-padded pressure. Therefore, in this method, high-wave number components of the pressure are averaged to estimate the noise variance, and the number of averaging components is predetermined. Moreover, the detailed explanation of the convergence of the extrapolated process is not described in the paper.

In this paper, the modified generalized cross-validation (GCV) method is applied for determination of the regularization parameter without a predetermined parameter. And, we describe the convergence of the extrapolation process in detail. As a demonstration example, experimental investigations are carried out to show the effectiveness of the proposed method.

II. DATA EXTRAPOLATION METHOD FOR BEM-BASED NAH

A. Review of BEM-based NAH

Before discussion of the data extrapolation method for BEM-based NAH, we review the basic theory of BEM-based NAH. The coordinate system $\mathbf{r}=(\xi_1,\xi_2,\xi_3)$ is shown in Fig. 1. The Helmholtz integral equation (HIE)[4] gives the pressure field $p(\mathbf{r})$ at a field point $\mathbf{r}$ in terms of the surface sound pressure $p(\mathbf{r}_s)$ and the normal component of the surface velocity $\nu_n(\mathbf{r})$ on the surface S_0

$$\beta p(\mathbf{r})=\int_{S_0}\left(i\rho_0 ckG(\mathbf{r}|\mathbf{r}_s)\nu_n(\mathbf{r}_s) - p(\mathbf{r}_s)\frac{\partial}{\partial n}G(\mathbf{r}|\mathbf{r}_s)\right)dS_0, \tag{1}$$

with

$$\beta=\begin{cases} 1 & \text{if } r \text{ is outside } S_0,\\ 1/2 & \text{if } r \text{ is on } S_0,\\ 0 & \text{if } r \text{ is inside } S_0,\end{cases} \tag{2}$$

where $G(\mathbf{r}|\mathbf{r}_s)=\exp(ik|\mathbf{r}-\mathbf{r}_s|)/(4\pi|\mathbf{r}-\mathbf{r}_s|)$ and $\partial/\partial n$ indicate the derivative with respect to the outward normal in the primed coordinate system. The sound speed is c, the fluid mass density is ρ_0, and $k=\omega/c=2\pi/\lambda$, with ω the frequency and λ the radiated wavelength.

Discretization of HIE is necessary to introduce BEM-based NAH. Here, the surface is broken up into N small elements of area, and the field point $\mathbf{r}$ is set on the surface. Then, Eq. (1) becomes

$$\frac{1}{2}p_j^S=\sum_{l=1}^{N}\left(p_l^S\frac{\partial \mathbf{G}_{jl}^s}{\partial n}-\mathbf{G}_{jl}^S\frac{\partial p_l^S}{\partial n}\right)\Delta S_l^S, \tag{3}$$

where the subscript refers to the small element number, the superscript S indicates that the field point is taken on the surface, p_j is the pressure, ΔS_l is the area of each element, and $\mathbf{G}_{jl}=[\exp(ik|\mathbf{r}_j-\mathbf{r}_l|)/(4\pi|\mathbf{r}_j-\mathbf{r}_l|)]$. Next, the HIE is discretized for the field point off the surface. Then, Eq. (1) becomes

$$p_j=\sum_{l=1}^{N}\left(p_l^S\frac{\partial \mathbf{G}_{jl}}{\partial n}-\mathbf{G}_{jl}\frac{\partial p_l^S}{\partial n}\right)\Delta S_l^S. \tag{4}$$

The vector, $\mathbf{p}^S=[p_j^S]$, can be eliminated from Eq. (4) by using Eq. (3) to yield the following single matrix equation for the pressure outside of the surface:

$$\mathbf{p}=\mathbf{H}\mathbf{v}_n, \tag{5}$$

where $\mathbf{v}_n$ is the normal velocity of surface, $\mathbf{H}$ is an $N\times N$ matrix, which is called a "vibro-acoustic transfer matrix."[14] Equation (5) can be inverted by applying the inversion matrix $\mathbf{H}^{-1}$

$$\mathbf{v}_n=\mathbf{H}^{-1}\mathbf{p}; \tag{6}$$

therefore, the surface velocity $\mathbf{v}_n$ can be reconstructed by Eq. (6) from the field pressure $\mathbf{p}$, which is measured on the concentric hologram surface. This is the basic concept of BEM-based NAH.

For actual implementation of BEM-based NAH, an appropriate regularization is necessary, because the inversion of $\mathbf{H}$ is very ill-conditioned due to the existence of evanescent-like waves. The regularization is implemented by singular value decomposition (SVD) and regularization filter. The SVD applied to $\mathbf{H}$ in Eq. (6) is

$$\mathbf{H}=\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{7}$$

where

$$\begin{aligned}&\mathbf{\Sigma}=\mathrm{diag}(\sigma_1,\sigma_2,\ldots,\sigma_N),\\ &\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_N\geqslant 0,\\ &\mathbf{U}\mathbf{U}^H=\mathbf{V}\mathbf{V}^H=\mathbf{I},\end{aligned} \tag{8}$$

where $\mathbf{U}$ and $\mathbf{V}$ mean the vectors, each of which has orthonormal columns, $\mathbf{I}$ is the identity matrix, and σ_i is the singular value. The column vectors $\{u_i\}$ and $\{v_i\}$ mean the wave vectors that decompose the distribution of field pressure on the hologram and that of normal velocity on the structure surface at a selected frequency. So, the inverse of the transfer matrix $\mathbf{H}^{-1}$ is

$$\mathbf{H}^{-1}=(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H)^{-1}=\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H, \tag{9}$$

where

$$\mathbf{\Sigma}^{-1}=\mathrm{diag}(1/\sigma_1,1/\sigma_2,\ldots,1/\sigma_N). \tag{10}$$

In the reconstruction process of surface velocity, high-order components of singular values affect the reconstructed result dominantly and they are correlated with the evanescent wave components. Usually the measured pressure is contaminated with the noise. Then, the measurement noise is amplified greatly by the high-order component, σ_i. Therefore, an appropriate regularization filter is necessary to suppress the amplification of the noise. Some regularization procedures have been proposed,[10,12,16–23] and the appropriate regularization procedure of NAH has been proposed by Williams.[24] His method is based on the improved Tikhonov regularization procedure. In his method, the regularization parameter is determined by the modified generalized cross-validation (GCV) method, which does not require *a priori* knowledge of the noise variance.

B. Influence of the finiteness of the measurement aperture

Maynard and Williams[1,4] have discussed the finite measurement aperture effect of FFT-based NAH in detail. There are two problems induced by the finiteness of the measurement aperture. One is ''wraparound error,'' and the other is ''discontinuity of the aperture edge.'' This wraparound error is inherent in FFT-based NAH. Therefore, BEM-based NAH does not suffer from the wraparound error. However, the discontinuity of the aperture edge causes a contamination of reconstructed result in BEM-based NAH. In this subsection, the influence of the discontinuity of the aperture edge is explained.

The discontinuity of the aperture edge induces two difficulties. One is the influence of ''error pressure,'' and the other is ''deformation of K-space spectrum.'' First, let us explain the influence of error pressure. The coordinate system $\mathbf{r}=(\xi_1,\xi_2,\xi_3)$ is shown in Fig. 1. Here, the finite aperture sets as S_w on the measurement surface S_m. The number of the discretized area on the complete measurement surface is set as N and that on the measurement aperture S_w is set as $n(N>n)$. The zero-padded measured pressure $\mathbf{p}_{\mathrm{m}}^{\delta}$ on the measurement surface S_m (abbreviated as ''restricted pressure,'' hereafter) is expressed as

$$\mathbf{p}_{\mathrm{m}}^{\delta}=\mathbf{W}\mathbf{p}^{\delta}=\mathbf{p}^{\delta}+(\mathbf{W}-\mathbf{I})\mathbf{p}^{\delta}=\mathbf{p}^{\delta}+\mathbf{p}_{\mathrm{error}}^{\delta}, \tag{11}$$

$$\mathbf{W}\equiv\mathrm{diag}\{W_j\}, \tag{12}$$

$$W_j=\begin{cases}1, & \mathbf{r}_j\in S_w\\ 0, & \mathbf{r}_j\notin S_w\end{cases}, \tag{13}$$

where the superscript δ indicates a quantity with noise. The normal velocity $\mathbf{v}_n^{\delta}$ is obtained from the pressure $\mathbf{p}^{\delta}$ measured on the complete measurement surface (abbreviated as ''full aperture pressure,'' hereafter) by BEM-based NAH with appropriate regularization filter. The error of the reconstructed result is small. However, the pressure $\mathbf{p}_{\mathrm{error}}^{\delta}$ contaminates the reconstructed result. This pressure is called error pressure. In this paper, we attempt to suppress the influence of this error pressure by the K-space data-extrapolation method. The error pressure is zero on the measurement aperture. And, the influence of the error pressure is large near the rim of the measurement aperture by the inherent property of the Green function that constitutes the vibro-acoustic transfer matrix.

Next, we examine the influence of the deformation of the K-space spectrum. To apply the K-space data extrapolation method, an appropriate band-limitation filter is necessary because K-space data extrapolation is based on the concept of the ''error norm reduction method.'' [25] That is, a band limitation of the pressure of the measurement surface is necessary to reduce the error norm. This constraint of the band limitation depends on the distribution of the K-space spectrum of the pressure. When the measurement pressure is contaminated by noise, an appropriate regularization filter is applicable as this band-limitation filter for the K-space data extrapolation method. In this paper, improved Tikhonov regularization filter (abbreviated as ''Tikhonov filter,'' hereafter) is used as this band-limitation filter.[24] The appropriate regularization parameter for Tikhonov filter is determined from the K-space spectrum of the full aperture pressure $\mathbf{U}^{\mathrm{H}}\mathbf{p}^{\delta}$ by the modified GCV method. However, the K-space spectrum of the restricted pressure is expressed as

$$\mathbf{U}^{\mathrm{H}}\mathbf{p}_{\mathrm{m}}^{\delta}=\mathbf{U}^{\mathrm{H}}\mathbf{p}^{\delta}+\mathbf{U}^{\mathrm{H}}\mathbf{p}_{\mathrm{error}}^{\delta}=\mathbf{U}^{\mathrm{H}}\mathbf{p}^{\delta}+\mathbf{U}^{\mathrm{H}}(\mathbf{W}-\mathbf{I})\mathbf{p}^{\delta}. \tag{14}$$

Here, the rank of the vectors $\mathbf{U}^{\mathrm{H}}$ is N, and each column vector is linearly independent. On the contrary, the rank of the vectors $\mathbf{U}^{\mathrm{H}}(\mathbf{W}-\mathbf{I})$ is $(N-n)$. And, each column vector is linearly dependent. Therefore, the K-space spectrum of the error pressure is deformed. This deformation of the spectrum prevents us from estimating the appropriate regularization parameter by the modified GCV method, and the Tikhonov filter cannot sufficiently suppress the amplification of high-order components of the error pressure and noise. As a result, the reconstructed result is seriously contaminated. In the next section, the influences of error pressure and deformation of K-space spectrum are examined by experimental results in detail.

C. Algorithm of data extrapolation method for BEM-based NAH

In this subsection, the algorithm of data extrapolation method for BEM-based NAH is briefly described. This method is the extension of the K-space data extrapolation method for FFT-based NAH.[15] As mentioned above, the appropriate regularization filter is necessary for applying the K-space data extrapolation method. However, this filter cannot be directly estimated from the restricted pressure. Therefore, the proposed method consists of two parts. One is a data extrapolation procedure based on the K-space data extrapolation method, whose cutoff wave number is fixed. The other is an estimation procedure of an appropriate regularization filter.

First, we explain the data extrapolation procedure. Next, an estimation procedure of an appropriate regularization filter is explained. Let us assume that an appropriate regularization filter is known. Here, the appropriate regularization filter effectively eliminates the noise components without changing the pressure $\mathbf{p}$; therefore, the cutoff wave number

of this filter is higher than the highest wave number of the pressure $\mathbf{p}$.[24] The Tikhonov filter,[24] which is used as the band-limitation filter for the K-space data extrapolation method, is expressed as

$$F^{\alpha,1} \equiv \mathrm{diag}\left(\cdots \frac{|\sigma_i|^2}{|\sigma_i|^2 + \alpha[\alpha/(\alpha+|\sigma_i|^2)]^2} \cdots \right). \tag{15}$$

The regularization parameter α is determined by seeking to minimize the function

$$J^1(\alpha) = \frac{\|(\mathbf{I}-F^{\alpha,1})\mathbf{U}^{\mathrm{H}}\mathbf{p}^{\delta}\|^2}{[\mathrm{Tr}(\mathbf{I}-F^{\alpha,1})]^2}. \tag{16}$$

Then, the procedure of data extrapolation is as follows:[15]

Step (1) The measurement aperture size is virtually enlarged by zero padding.

Step (2) Smoothed pressure $\mathbf{p}_i^{\delta}$ is obtained as

$$\mathbf{p}_i^{\delta} = \mathbf{U}F^{\alpha,1}\mathbf{U}^{\mathrm{H}}\mathbf{p}_{i-1}'^{\delta}, \tag{17}$$

where i is the iteration number of step (2) to step (3), and $\mathbf{p}_0'^{\delta}$ is the restricted pressure $\mathbf{p}_{\mathrm{m}}^{\delta}$. If the iteration number exceeds the threshold, then go to step (4). The threshold of the iteration number is sufficiently large (in this paper, the iteration number is 1000).[15] Otherwise, go to step (3).

Step (3) The smoothed pressure is corrected according to the measured pressure

$$\mathbf{p}_i'^{\delta} = \mathbf{p}_{\mathrm{m}}^{\delta} + (\mathbf{I}-\mathbf{W})\mathbf{p}_i^{\delta}. \tag{18}$$

Then, return step (2).

Step (4) The surface velocity is reconstructed from the extrapolated pressure $\mathbf{p}_{1000}^{\delta}$.

The reduction of the error pressure $\mathbf{p}_{\mathrm{error}}^{\delta}$ is explained as follows. The restricted pressure $\mathbf{p}_{\mathrm{m}}^{\delta}$ is expressed as Eq. (11). The cutoff wave number of the filter $F^{\alpha,1}$ is higher than the highest wave number of the pressure $\mathbf{p}$; therefore, the pressure $\mathbf{p}$ cannot be modified by this data extrapolation process, and $\mathbf{p}^{\delta}$ is hardly varied in this process.[15] On the contrary, $\mathbf{p}_{\mathrm{error}}^{\delta}$ is modified by this process. The K-space spectrum of the error pressure is expressed as

$$\mathbf{U}^{\mathrm{H}}\mathbf{p}_{\mathrm{error}}^{\delta} = \mathbf{U}^{\mathrm{H}}(\mathbf{W}-\mathbf{I})\mathbf{p}^{\delta}. \tag{19}$$

The rank of the vectors $\mathbf{U}^{\mathrm{H}}(\mathbf{W}-\mathbf{I})$ is $(N-n)$, i.e., each column vector is linearly dependent. Therefore, lower-order wave number components of error pressure are dispersed over the high-order region. The band-limitation filter suppresses this leaked high-order wave number component. The smoothed error pressure

$$\mathbf{p}_{\mathrm{error},1}^{\delta} = \mathbf{U}F^{\alpha,1}\mathbf{U}^{\mathrm{H}}(\mathbf{W}-\mathbf{I})\mathbf{p}^{\delta}, \tag{20}$$

will leak into the measurement aperture. Correcting the smoothed error pressure in the measurement aperture to zero reduces the norm of the error pressure (step 3)

$$\mathbf{p}_{\mathrm{error},1}'^{\delta} = (\mathbf{I}-\mathbf{W})\mathbf{p}_{\mathrm{error},1}^{\delta}. \tag{21}$$

Thus, the error pressure is reduced in real and K spaces.[15]

Next, the estimation procedure of the appropriate regularization parameter α is explained. The appropriate regularization parameter of Tikhonov filter, α_{full}, can be determined from the full aperture pressure by the modified GCV method. However, the appropriate regularization parameter cannot be directly determined from the restricted pressure. Therefore, a regularization parameter $\alpha_{\mathrm{e},0}$ is set provisionally. Then, the data extrapolation procedure is applied with the Tikhonov filter to the restricted pressure. And, the estimated regularization parameter $\alpha_{\mathrm{e},1}$ is obtained from the extrapolated pressure by the modified GCV method. When the provisionally defined parameter $\alpha_{\mathrm{e},0}$ equals α_{full}, this estimated parameter $\alpha_{\mathrm{e},1}$ becomes α_{full}. The proposed procedure of the estimation of regularization parameter is based on this relationship, whose detail is explained as follows.

First, we explain the case where $\alpha_{\mathrm{e},0} < \alpha_{\mathrm{full}}$. In this case, as mentioned above, the full aperture pressure $\mathbf{p}^{\delta}$ is hardly modified in the data extrapolation process. On the other hand, the error pressure component $\mathbf{p}_{\mathrm{error}}^{\delta} = (\mathbf{I}-\mathbf{W})\mathbf{p}^{\delta}$ above the cutoff wave number is reduced by the extrapolation procedure. Therefore, the spectrum of the extrapolation pressure is concentrated in the low-wave number region. In other words, the wave number component of the extrapolated pressure above the predetermined cutoff wave number, which corresponds to the regularization parameter $\alpha_{\mathrm{e},0}$, is eliminated by the extrapolation process. As the result, the cutoff wave number estimated from the extrapolated pressure is lower than the predetermined cutoff wave number, and the corresponding regularization parameter $\alpha_{\mathrm{e},1}$ is larger than that $\alpha_{\mathrm{e},0}$.

Next, we consider the case that $\alpha_{\mathrm{e},0} > \alpha_{\mathrm{full}}$. From Eqs. (11) and (17), we see that the smoothed result of the restricted pressure by the Tikhonov filter is

$$\begin{aligned}\mathbf{p}_1^{\delta} &= \mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}_{\mathrm{m}}^{\delta} = \mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}^{\delta} + \mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}_{\mathrm{error}}^{\delta} \\ &= \mathbf{p}_{\alpha_{\mathrm{e},0},1}^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}^{\delta},\end{aligned} \tag{22}$$

where

$$\mathbf{p}_{\alpha_{\mathrm{e},0},1}^{\delta} = \mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}^{\delta}, \tag{23}$$

$$\mathbf{p}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}^{\delta} = \mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}p_{\mathrm{error}}^{\delta}. \tag{24}$$

The smoothed pressure is corrected according to the restricted pressure [step (3)]. To obtain the corrected pressure, we substitute Eq. (22) to Eq. (18), then

$$\begin{aligned}\mathbf{p}_1'^{\delta} &= \mathbf{p}_{\mathrm{m}}^{\delta} + (\mathbf{I}-\mathbf{W})(\mathbf{p}_{\alpha_{\mathrm{e},0},1}^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}^{\delta}) \\ &= \mathbf{p}^{\delta} + \mathbf{p}_{\mathrm{error}}^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},1}'^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}'^{\delta},\end{aligned} \tag{25}$$

where

$$\mathbf{p}_{\alpha_{\mathrm{e},0},1}'^{\delta} = (\mathbf{I}-\mathbf{W})\mathbf{p}_{\alpha_{\mathrm{e},0},1}^{\delta}, \tag{26}$$

$$\mathbf{p}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}'^{\delta} = (\mathbf{I}-\mathbf{W})\mathbf{p}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}^{\delta}. \tag{27}$$

The smoothed result of this corrected pressure $\mathbf{p}_1'^{\delta}$ by the Tikhonov filter is

$$\begin{aligned}\mathbf{p}_2^{\delta} &= \mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}_1'^{\delta} \\ &= \mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}(\mathbf{p}^{\delta} + \mathbf{p}_{\mathrm{error}}^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},1}'^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}'^{\delta}) \\ &= \mathbf{p}_{\alpha_{\mathrm{e},0},1}^{\delta} + \mathbf{p}_{\delta_{\mathrm{e},0},1,\mathrm{error}}^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},2}^{\delta} + \mathbf{p}_{\alpha_{\mathrm{e},0},2,\mathrm{error}}^{\delta},\end{aligned} \tag{28}$$

[step (2)], where

$$\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},2}=\mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},1}, \tag{29}$$

$$\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},2,\mathrm{error}}=\mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}. \tag{30}$$

This smoothed pressure is corrected by the restricted pressure, again [step (3)]. The corrected pressure is expressed as

$$\begin{aligned}\mathbf{p}'^{\delta}_{2}&=\mathbf{p}^{\delta}_{\mathrm{m}}+(\mathbf{I}-\mathbf{W})(\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},1}+\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}+\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},2}+\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},2,\mathrm{error}})\\&=\mathbf{p}^{\delta}+\mathbf{p}^{\delta}_{\mathrm{error}}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},1}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},2}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},2,\mathrm{error}},\end{aligned} \tag{31}$$

where

$$\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},2}=(\mathbf{I}-\mathbf{W})\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},2}, \tag{32}$$

$$\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},2,\mathrm{error}}=(\mathbf{I}-\mathbf{W})\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},2,\mathrm{error}}. \tag{33}$$

After l - time iterations of this process, the corrected pressure $\mathbf{p}'^{\delta}_{l}$ is expressed as

$$\mathbf{p}'^{\delta}_{l}=\mathbf{p}^{\delta}+\mathbf{p}^{\delta}_{\mathrm{error}}+\sum_{b=1}^{l}(\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b,\mathrm{error}}), \tag{34}$$

and

$$\mathbf{p}'^{\delta}_{\infty}=\lim_{l\to\infty}\mathbf{p}'^{\delta}_{l}=\mathbf{p}^{\delta}+\mathbf{p}^{\delta}_{\mathrm{error}}+\sum_{b=1}^{\infty}(\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b,\mathrm{error}}). \tag{35}$$

Here

$$\lim_{b\to\infty}\|\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b}\|=0, \tag{36}$$

$$\lim_{b\to\infty}\|\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b,\mathrm{error}}\|=0, \tag{37}$$

because the norms of $\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b}$ and $\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b,\mathrm{error}}$ are reduced in real and K space in the iteration process.[15] However

$$\|\mathbf{p}^{\delta}_{\mathrm{error}}+\sum_{b=1}^{\infty}(\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b,},\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},b,\mathrm{error}})\|\neq 0, \tag{38}$$

because $\alpha_{\mathrm{e},0}>\alpha_{\mathrm{full}}$. Since full aperture pressure is varied by the Tikhonov filter, the term $\mathbf{p}^{\delta}_{\mathrm{error}}=(\mathbf{W}-\mathbf{I})\mathbf{p}^{\delta}$ is not eliminated. The influence of the high-order wave number component of the pressure spectrum is dominant in the estimation process of the regularization parameter by the modified GCV method.[24] This error pressure term includes a spurious high-order wave number component. Therefore, the regularization parameter $\alpha_{\mathrm{e},1}$, which is estimated from the extrapolated pressure $\mathbf{p}'^{\delta}_{\infty}$ by the modified GCV method, is smaller than the appropriate parameter. In this way, if $\alpha_{\mathrm{e},0}$ is larger than α_{full}, then $\alpha_{\mathrm{e},1}$ is smaller than $\alpha_{\mathrm{e},0}$.

Finally, we examine the case that $\alpha_{\mathrm{e},0}=\alpha_{\mathrm{full}}$. In this case, as mentioned above, the appropriate regularization filter is applied; therefore, the full pressure term $\mathbf{p}^{\delta}$ in Eq. (11) is hardly modified in the data extrapolation process. On the contrary, the error pressure term $\mathbf{p}^{\delta}_{\mathrm{error}}$ is modified. Therefore, the corrected pressure $\mathbf{p}'^{\delta}_{1}$ is expressed as

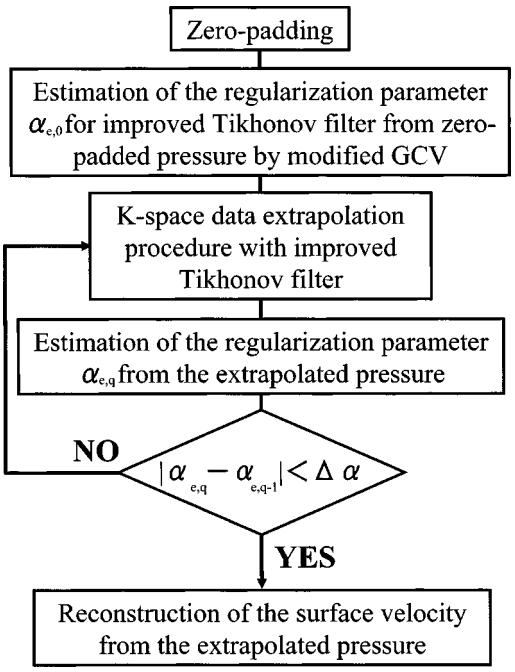


FIG. 2. Procedure of the proposed data extrapolation method.

$$\begin{aligned}\mathbf{p}'^{\delta}_{1}&=\mathbf{p}^{\delta}_{\mathrm{m}}+(\mathbf{I}-\mathbf{W})(\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},1}+\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},1,\mathrm{error}})\\&\approx\mathbf{W}\mathbf{p}^{\delta}+(\mathbf{I}-\mathbf{W})\mathbf{p}^{\delta}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},1,\mathrm{error}}\\&=\mathbf{p}^{\delta}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},1,\mathrm{error}},\end{aligned} \tag{39}$$

where

$$\mathbf{p}^{\delta}_{\alpha_{\mathrm{e},0},1}=\mathbf{U}F^{\alpha_{\mathrm{e},0},1}\mathbf{U}^{\mathrm{H}}\mathbf{p}^{\delta}\approx\mathbf{p}^{\delta}, \tag{40}$$

because $\alpha_{\mathrm{e},0}=\alpha_{\mathrm{full}}$. After l - time iterations of this process, the corrected pressure $\mathbf{p}'^{\delta}_{l}$ is expressed as

$$\mathbf{p}'^{\delta}_{l}\approx\mathbf{p}^{\delta}+\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},l,\mathrm{error}}, \tag{41}$$

and

$$\mathbf{p}'^{\delta}_{\infty}=\lim_{l\to\infty}\mathbf{p}'^{\delta}_{\infty}=\mathbf{p}^{\delta}+\lim_{l\to\infty}\mathbf{p}'^{\delta}_{\alpha_{\mathrm{e},0},l,\mathrm{error}}=\mathbf{p}^{\delta} \tag{42}$$

[cf. Eq. (37)]. Therefore, the spurious high-order wave number component is diminished by data extrapolation procedure, and the estimated regularization parameter $\alpha_{\mathrm{e},1}$ from the extrapolated pressure $\mathbf{p}'^{\delta}_{\infty}$ by the modified GCV method equals α_{full}.

The above-mentioned discussion is summarized as

$$\alpha_{\mathrm{e},1}\begin{cases}<\alpha_{\mathrm{e},0},\\=\alpha_{\mathrm{e},0},\\>\alpha_{\mathrm{e},0},\end{cases}\quad\text{when}\quad\begin{cases}\alpha_{\mathrm{full}}<\alpha_{\mathrm{e},0}\\\alpha_{\mathrm{full}}=\alpha_{\mathrm{e},0}.\\\alpha_{\mathrm{full}}>\alpha_{\mathrm{e},0}\end{cases} \tag{43}$$

This relationship is always satisfied in every estimation procedure of the regularization parameter. Therefore, this procedure enables us to estimate the appropriate regularization parameter from the extrapolated pressure.

The algorithm of the proposed data extrapolation method for BEM-based NAH is shown in Fig. 2. A brief explanation is as follows.

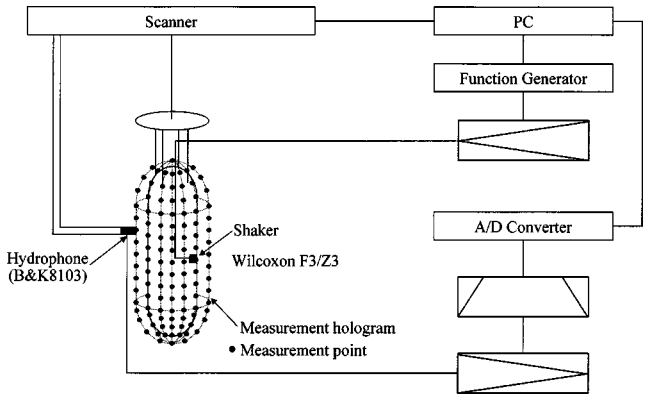


FIG. 3. A block diagram of the measurement system for the BEM-based NAH experiment.

Step (1) The measurement aperture size is virtually enlarged by zero padding.

Step (2) Initial regularization parameter $\alpha_{e,0}$ is estimated by the modified GCV method from the zero-padded pressure, and the band-limitation filter for data extrapolation procedure is set to be $F^{\alpha_{e,0},1}$.

Step (3) Using the band-limitation filter, the data extrapolation procedure, which is explained above, is applied to the virtually enlarged pressure.

Step (4) The regularization parameter $\alpha_{e,q}$ is estimated from the extrapolated pressure by the modified GCV method. Here, q is the iteration number of step (3) to step (5).

Step (5) If $|\alpha_{e,q}-\alpha_{e,q^{-1}}|<\Delta\alpha$, the surface velocity is reconstructed from the extrapolated pressure. The regularization filter is set as $F^{\alpha_{e,q},1}$. And, this extrapolation procedure is terminated. Otherwise, the band-limitation filter is set to be $F^{\alpha_{e,q},1}$. Then, go to step (3). In this study, $\Delta\alpha=\alpha_{e,0}\times10^{-4}$.

III. MODEL EXPERIMENT

To examine the effectiveness of the proposed method, the distribution of the surface velocity of a submerged shell with spheroidal caps was reconstructed from the measured pressure hologram by the proposed method. The model used in the experiment is aluminum shell with spherical caps (length is 800 mm, radius is 100 mm).

The experiment was carried out in the 5th Research Center's water pool, 15 m long, 9 m wide, and 8 m deep. An automated scanning system was developed for the experiment at this facility (Fig. 3). A sample shell is immersed at about 4 m depth. A shaker (Wilcoxon F3/83) that vibrates the sample shell is attached at the center of the shell. The shaker is driven by a broadband (300–6000 Hz) LFM signal. The full measurement aperture (that is, the location of a hydrophone's [Brüel & Kjaer type 8103] acoustic center) is set near the sample shell. The distance between the full measurement aperture and the sample shell is 30 mm. The full measurement aperture is scanned by stepwise horizontal and rotational movement with a robotic scanner. The measured data of the full hologram surface are taken in 25-mm steps along the vertical axis and in 11.25-deg steps along the rotational axis. Therefore, 32×32 data points make up the full measurement aperture. The mesh size is small compared with the acoustic wavelength of interest.

The received and amplified signal from the hydrophone is digitized by an Agilent E1437A A/D converter. The digitized signal is transferred to a PC through IEEE 1394 bas. Data sampling rate is 1.25 MHz. Data sampling inception time is precisely controlled by a PC, which acts as the reference source for holographic interference. All postexperimental data processing is executed on a PC.

First, let us examine the dependence of the Tikhonov filter on the SNR. Figure 4 shows the reconstructed surface velocity from the full aperture pressure (SNR=20 dB) and Tikhonov filter. In this paper, the reconstructed results are normalized by the maximum value. The driving frequency is 1543 Hz. Figure 4(a) shows the three half-length modes along the vertical axis and two full-length modes along the rotational axis. Figure 4(b) shows the regularization filter. In this study, random Gaussian noise was added to the hologram with predetermined variance to vary the SNR. According to the deterioration of the SNR, the cutoff singular value becomes lower. In the results which follow, the ''exact'' velocity on the shell were determined by a standard BEM-based NAH reconstruction[4,24] using the full measurement aperture pressure (at 1543 Hz, SNR=50 dB). The error of the reconstructed surface velocity is defined as follows:

$$E=\frac{\|\mathbf{W}(\mathbf{v}_{n,\mathrm{exact}}-\mathbf{v}_{n,\mathrm{reconst}})\|}{\|\mathbf{W}\mathbf{v}_{n,\mathrm{exact}}\|}, \tag{44}$$

where $\mathbf{v}_{n,\mathrm{exact}}$ is the exact velocity on the shell, $\mathbf{v}_{n,\mathrm{reconst}}$ is the surface velocity reconstructed, and $\mathbf{W}$ is the measurement aperture. (When the surface velocity is reconstructed from the full aperture pressure, this measurement aperture corresponds to full aperture.) The error when reconstructed from the full aperture pressure of SNR=20 dB is 25.4%.

Now, the influence of the deformation of the K-space spectrum is examined. Figure 5 shows the K-space spectrum of the pressure on the measurement hologram, which is normalized by the maximum value (SNR=20 dB), and Tikhonov filter. The black line of Fig. 5(a) is the spectrum of the full aperture pressure. The amplitude of the high-order wave number component of the full aperture pressure is relatively small. Therefore, the cutoff singular value of the corresponding Tikhonov filter is relatively small [Fig. 5(b)]. The red line is the spectrum of the restricted pressure. The measurement aperture S_W is $-400<z<400$ mm in vertical axis and $22.5<\theta<90$ deg in rotational axis [cf. Fig. 7(a)]. The spectrum of the restricted pressure is seriously deformed, and the amplitude of high-order wave number components of the restricted pressure is enormously large compared with that of the full aperture spectrum. Therefore, the cutoff singular value of the corresponding Tikhonov filter is high, and this filter cannot sufficiently suppress the amplification of high-order wave number component of the error pressure and noise. Due to this, the reconstructed result is seriously contaminated by the error pressure and noise. The blue line is the spectrum of the extrapolated pressure from the restricted pressure by the proposed method. Comparison between the spectrum of the extrapolated pressure and that of the full

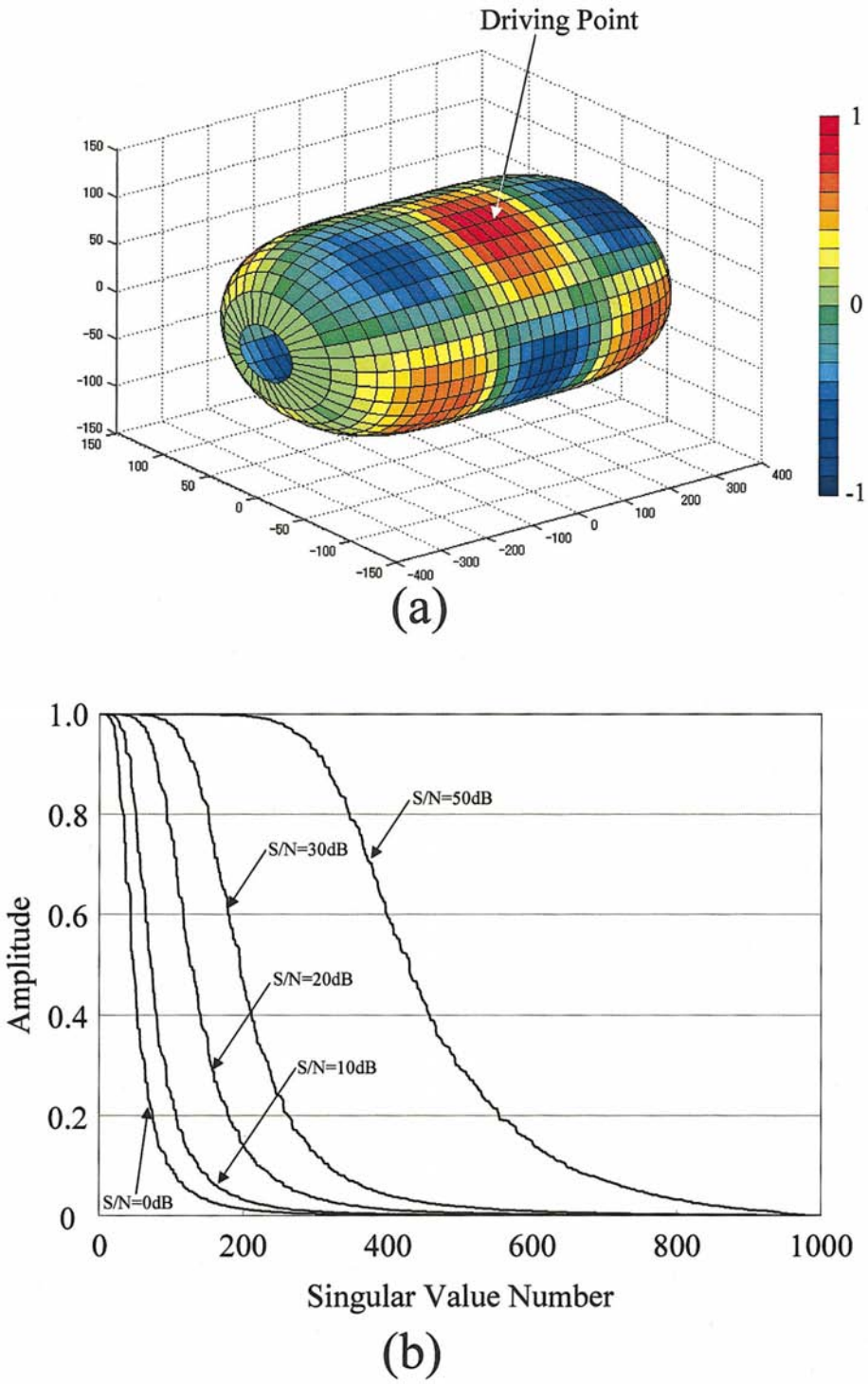
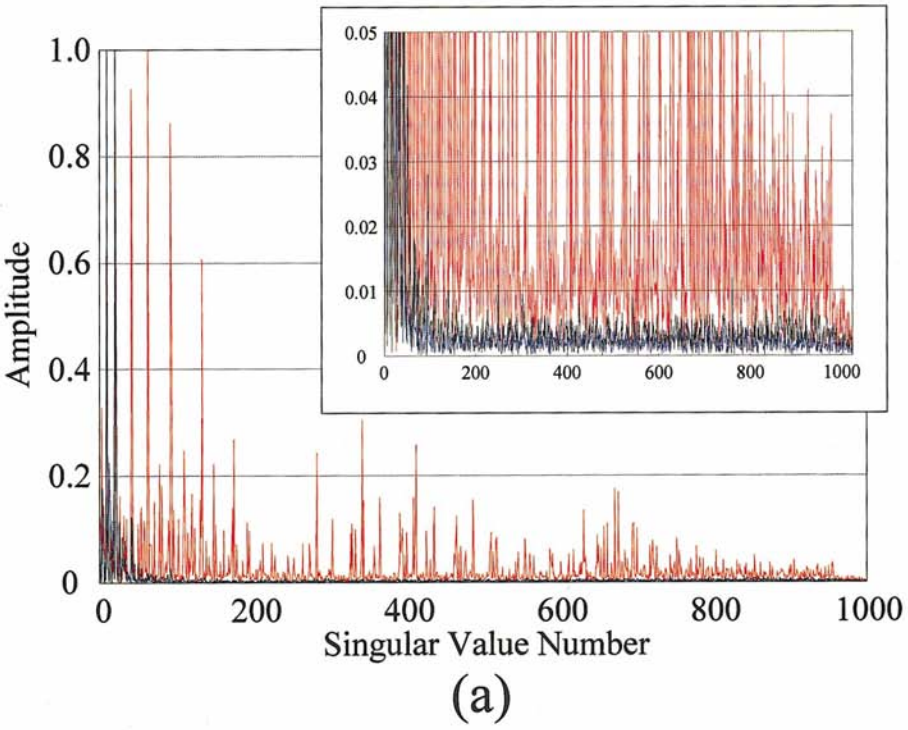


FIG. 4. Surface velocity reconstructed from the full-aperture pressure and Tikhonov filter. The driving frequency is 1543 Hz. (a) The real part of the surface velocity reconstructed from the full-aperture pressure. SNR $=20$ dB. (b) Filter shapes of the improved Tikhonov filters. The SNRs are changed from 0 to 50 dB.

aperture pressure shows a good agreement. Therefore, the filter shape of the Tikhonov filter estimated from the extrapolated pressure is similar to that estimated from the full aperture pressure. And, the amplitude of high-order wave number components of this extrapolated pressure is sufficiently suppressed by this filter. This result infers the reliability of the proposed method.

The error pressure and reconstructed surface velocity from the error pressure are shown in Fig. 6 to clarify the influence of the error pressure. Figure 6(a) is the error pressure without extrapolation. The flat area corresponds to the measurement aperture. Precipitous discontinuity exists at the rim of the measurement aperture. And, the amplitude near the measurement aperture is relatively large. Figure 6(b) displays the reconstructed surface velocity from the error pressure without extrapolation on the measurement aperture. This surface velocity corresponds to the reconstructed error. Figure 6(c) is the error pressure after extrapolation procedure by the proposed method. The discontinuity at the rim of the measurement aperture is sufficiently smoothed, and the error pressure near the measurement aperture is suppressed. We note that the data extrapolation process pushes the rim discontinuity toward the external area while keeping the measured data less affected.[15] Figure 6(d) is the reconstructed



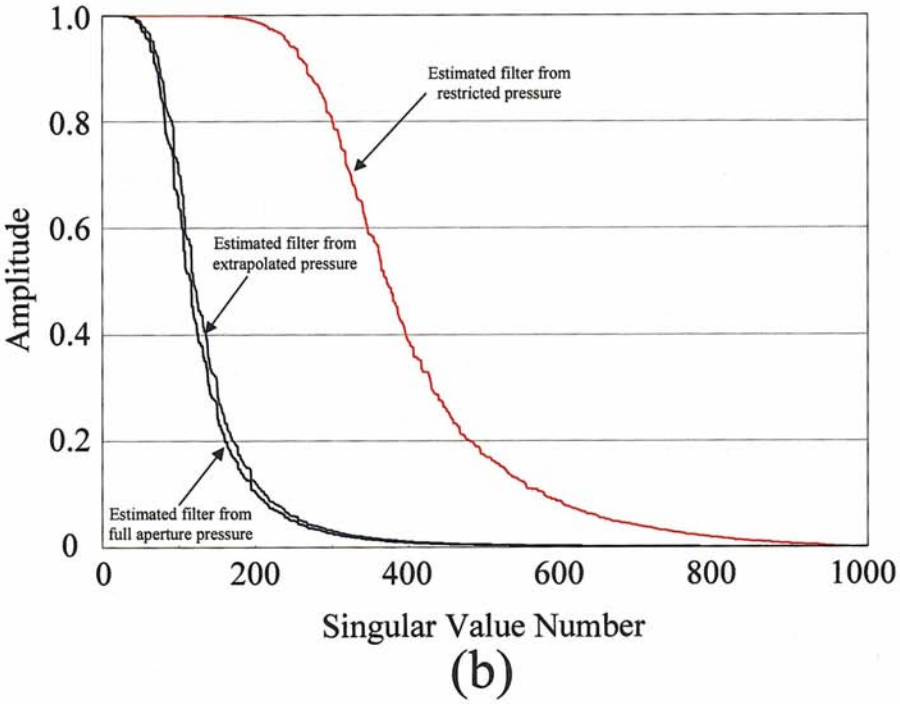


FIG. 5. The influence of the deformation of K-space spectrum upon the filter shape of the improved Tikhonov filter. The frequency is 1543 Hz and SNR = 20 dB. (a) K-space spectra of the pressure on the measurement hologram. These spectra are normalized by the maximum values, respectively. The black and red lines are the spectrums of the full-aperture pressure and that of the restricted pressure, respectively. The blue line indicates the spectrums of the extrapolated pressure by the proposed method. (b) Filter shapes of the improved Tikhonov filters.

surface velocity from the error pressure after the extrapolation procedure. Comparison between Figs. 6(b) and (d) shows that the amplitude of the reconstructed error is suppressed by data extrapolation. This result exhibits the effectiveness of the proposed method.

Figure 7 illustrates the reconstructed result of the surface velocity. The driving frequency is 1543 Hz. Figure 7(a) shows the surface velocity on the measurement aperture, which is reconstructed from the full-aperture pressure data (SNR = 20 dB). The reconstruction error on the measurement aperture is 25.9%. The surface velocity reconstructed from the restricted pressure is shown in Fig. 7(b). The reconstruction error is 93%. The comparison between Figs. 7(a) and (b) indicates that the distribution of the surface velocity is seriously deformed by the finiteness of the measurement aperture. Figure 7(c) is the surface velocity reconstructed from the pressure that is obtained by the proposed method. The reconstruction error on the measurement aperture is 29.5%. The difference between the reconstruction error from the full-aperture pressure and that from the extrapolated pres-

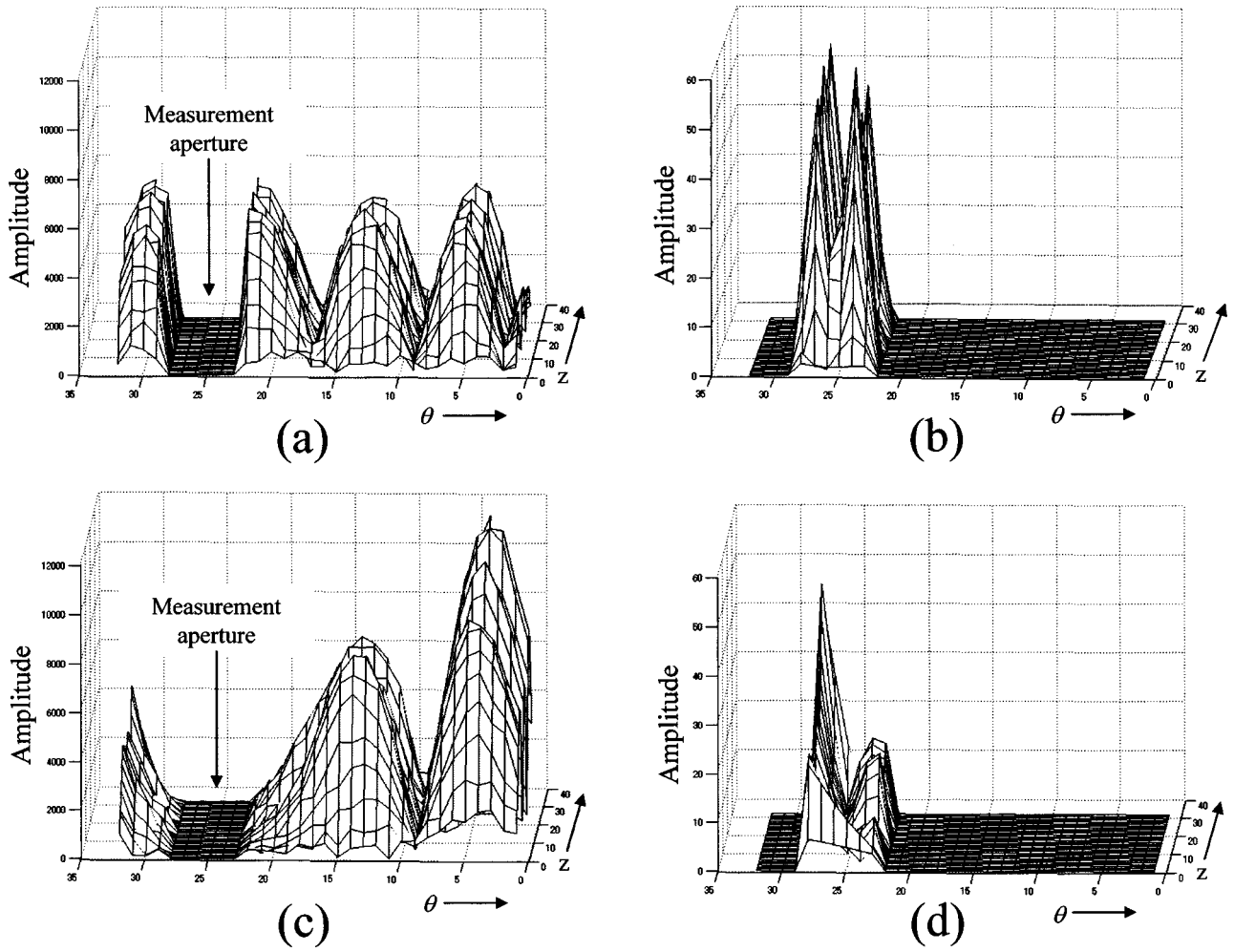


FIG. 6. The influence of the error pressure upon the reconstructed result. The frequency is 1543 Hz and SNR = 20 dB. (a) The amplitude of the error pressure on the measurement hologram. (b) The amplitude of the surface velocity reconstructed from the error pressure on the measurement aperture. (c) The amplitude of the error pressure on the measurement hologram after the extrapolation procedure. (d) The amplitude of the surface velocity reconstructed from the error pressure after extrapolation procedure.

sure is 3.6%, and Fig. 7(c) seems to show excellent agreement with Fig. 7(a). The relationship between the regularization parameter of the Tikhonov filter and the error of the reconstructed velocity is shown in Fig. 7(d). The regularization parameter estimated from the restricted pressure by the modified GCV method (α_{restrict}) is too small. And, the reconstruction error from the restricted pressure (rectangle) is very large. The reconstructed error from the pressure that is extrapolated with the Tikhonov filter $F^{\alpha_{\text{restrict}},1}$ (diamond) is 75%. Despite applying the data extrapolation procedure, the reconstruction error is large because the estimated regularization parameter is too small. Therefore, the estimation of an appropriate regularization parameter for data extrapolation is crucial for reducing the reconstruction error. On the other hand, the reconstruction error from the extrapolated pressure (solid line) and the regularization parameter estimated from the proposed method (circle) show that the proposed method provides good result. The regularization parameter estimated from the full-aperture pressure by the modified GCV method α_{full} (vertical broken line) and that estimated from the restricted pressure by the proposed method α_{propose} (circle) show a good agreement. As a result, the propriety of Eq. (43) is exhibited.

Figure 8 is the reconstructed result of the surface velocity. The driving frequency is 4101.6 Hz. In this frequency, the ''exact'' velocity on the shell was obtained from the full measurement aperture pressure (SNR = 30 dB). Figure 8(a) is the surface velocity reconstructed from the full-aperture pressure data (SNR = 20 dB). The reconstruction error on the measurement aperture is 14.9%. Figure 8(b) is the surface velocity reconstructed from the restricted pressure. The reconstruction error is 127%, and the reconstructed result is seriously deformed by the finiteness of the measurement aperture. The surface velocity reconstructed from the pressure extrapolated by the proposed method is shown in Fig. 8(c). The reconstruction error is 22.9%. The difference between the reconstruction error from the full-aperture pressure and that from the extrapolated pressure is 8%, and Fig. 8(c) seems to show good agreement with Fig. 8(a). The relationship between the regularization parameter and the reconstruction error is shown in Fig. 8(d). The reconstruction error from the extrapolated pressure and the regularization parameter estimated from the proposed method show good result. And, the regularization parameter estimated from the restricted pressure by the proposed method shows good agreement with that estimated from the full-aperture pressure by

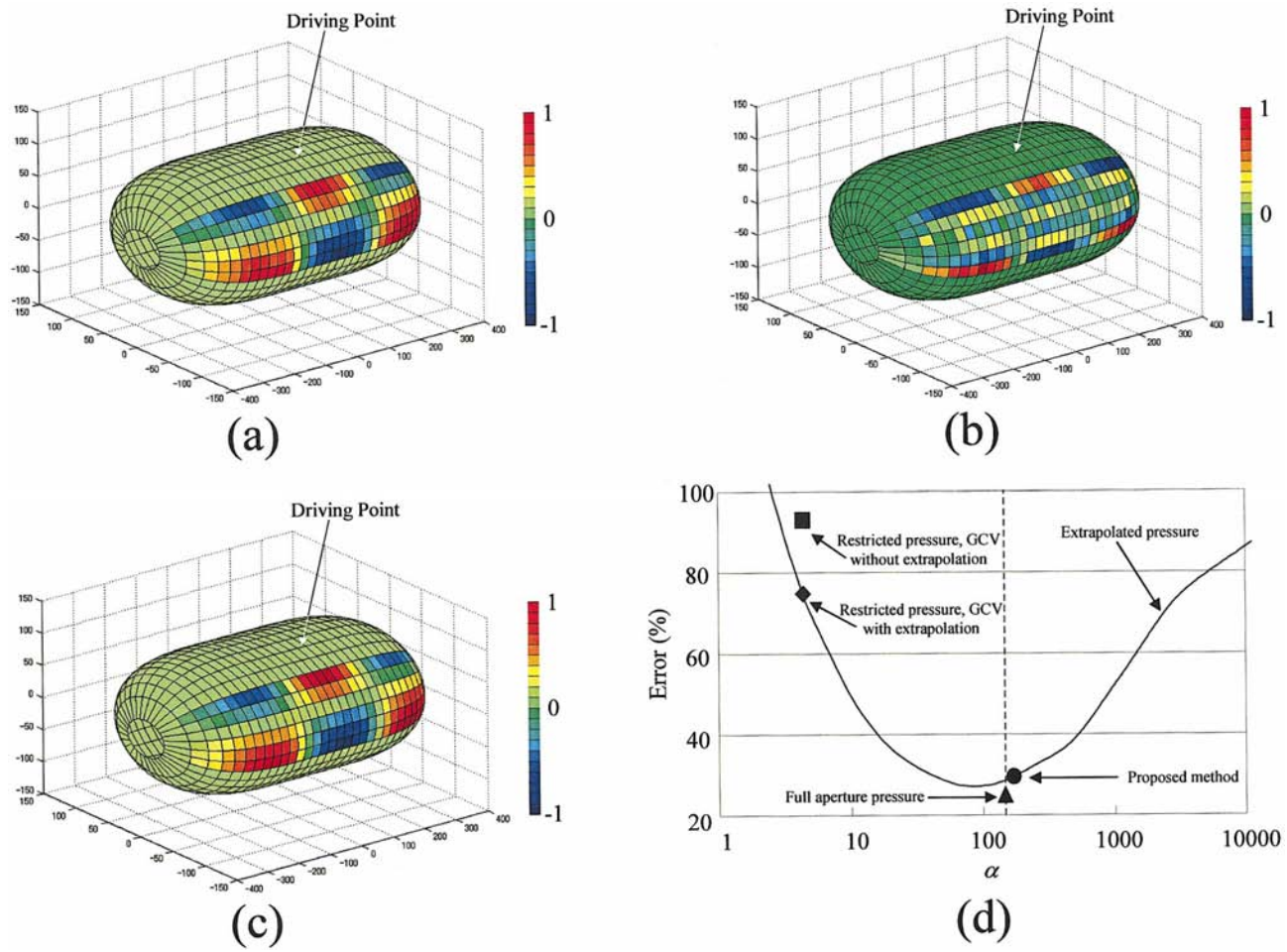


FIG. 7. Reconstruction result of surface velocity. The structure is driven at 1543 Hz and SNR = 20 dB. The reconstructed results are shown on the measurement aperture. (a) Surface velocity reconstructed from the full-aperture pressure data. The regularization parameter is estimated from the full-aperture pressure by the modified GCV method. (b) Surface velocity reconstructed from the restricted pressure. The regularization parameter is estimated from the restricted pressure by modified GCV method. (c) Surface velocity reconstructed from the restricted pressure by the proposed method. (d) The relationship between regularization parameters and reconstruction errors. The solid line is the reconstruction error from the extrapolated pressure. The vertical broken line indicates the regularization parameter estimated from the full-aperture pressure by the modified GCV method. The rectangle and diamond are the reconstruction errors from the restricted pressure without extrapolation and that from the restricted pressure with extrapolation, respectively. The circle and rectangle are the reconstruction error from the restricted pressure by the proposed method and that from the full-aperture pressure, respectively.

the modified GCV method. On the other hand, the regularization parameter estimated from the restricted pressure by the modified GCV method α_{restrict} is too small. The reconstruction error from the restricted pressure (127%) and that from the extrapolated pressure with the Tikhonov filter $F^{\alpha_{\text{restrict}},1}$ (98%) are very large.

As a result, if the measurement aperture is small compared with the structure and the finiteness of the measurement aperture should be considered, the K-space data extrapolation method is recommended to reconstruct the surface velocity. However, a regularization parameter of band-limitation filter, which is estimated from the restricted pressure by the modified GCV method, is always smaller than the appropriate value. Therefore, the reconstructed result is seriously contaminated by the spurious high-order wave number component and noise, and reconstruction error becomes large.

On the contrary, the proposed method enables us to estimate the appropriate regularization parameter from the restricted pressure, and the influences of the finiteness of the measurement aperture are sufficiently suppressed. Therefore, the reconstruction error of the proposed method becomes small.

IV. CONCLUSIONS

A new data extrapolation method for BEM-based NAH has been presented. The finiteness of the measurement aperture precipitates the influence of ''error pressure.'' Therefore, the reconstructed result is seriously contaminated by noise and spurious high-order wave number components. Moreover, ''deformation of K-space spectrum'' is also induced, and an appropriate regularization filter cannot be directly estimated from the restricted pressure. To overcome this situation, a new data extrapolation method is proposed. This method is an extension of the K-space data extrapolation method for FFT-based NAH. This method includes an estimation procedure of the appropriate regularization parameter. In this procedure, the regularization parameter of the band-limitation filter for data extrapolation method is automatically determined by modified GCV method from the restricted pressure.

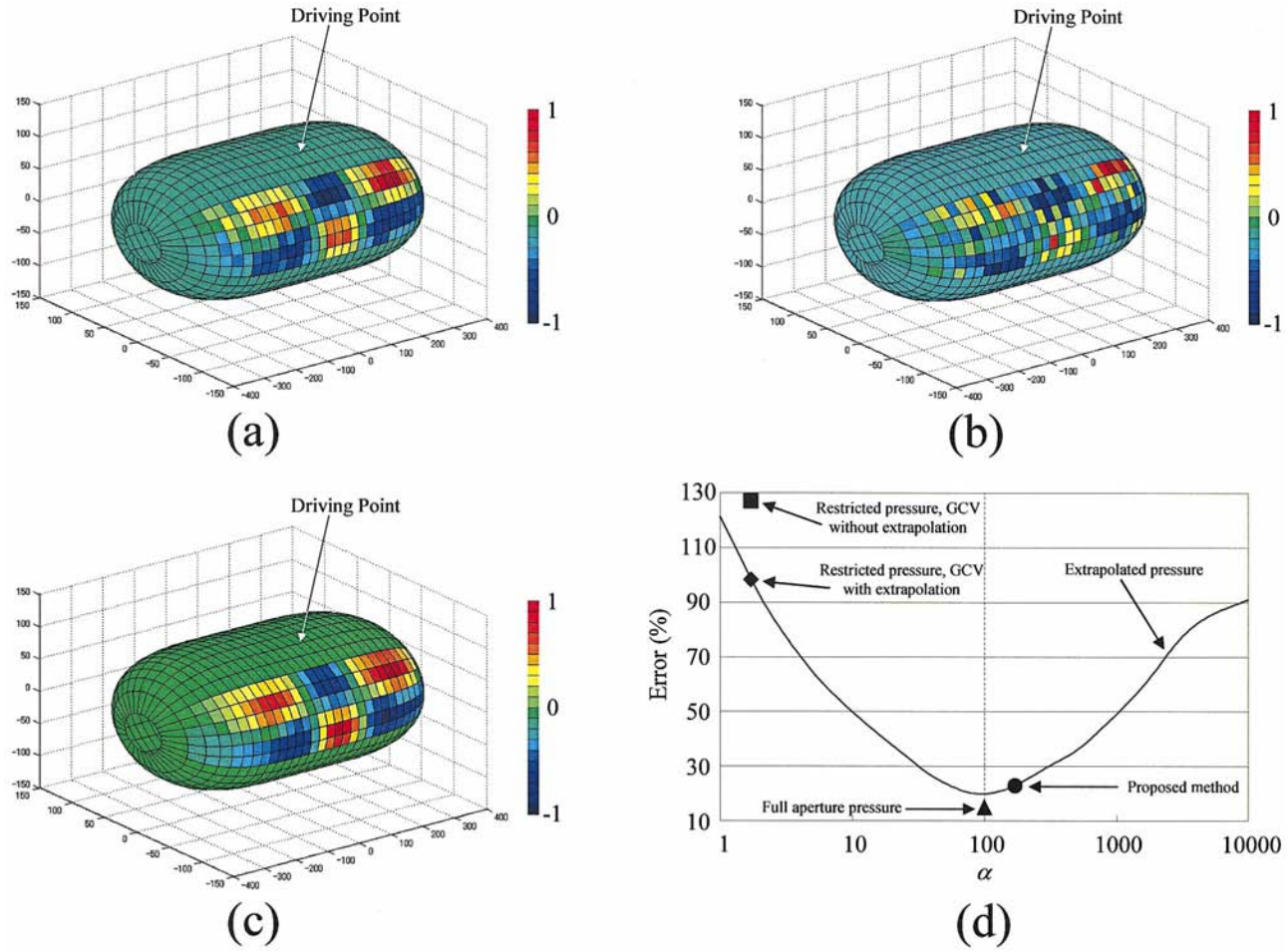


FIG. 8. Reconstruction result of surface velocity. The structure is driven at 4101.6 Hz and SNR = 20 dB. The reconstructed results are shown on the measurement aperture. (a) Surface velocity reconstructed from the full-aperture pressure data. The regularization parameter is estimated from the full-aperture pressure by the modified GCV method. (b) Surface velocity reconstructed from the restricted pressure. The regularization parameter is estimated from the restricted pressure by modified GCV method. (c) Surface velocity reconstructed from the restricted pressure by the proposed method. (d) The relationship between regularization parameters and reconstruction errors. The solid line is the reconstruction error from the extrapolated pressure. The vertical broken line indicates the regularization parameter estimated from the full-aperture pressure by the modified GCV method. The rectangle and diamond are the reconstruction errors from the restricted pressure without extrapolation and that from the restricted pressure with extrapolation, respectively. The circle and rectangle are the reconstruction error from the restricted pressure by the proposed method and that from the full-aperture pressure, respectively.

The effectiveness of the proposed method is confirmed by the experiment. For comparison, the reconstructed result of BEM-based NAH from the restricted pressure is obtained. The regularization parameter is estimated from the restricted pressure by the modified GCV method. The estimated regularization parameter from the restricted pressure by the modified GCV method is too small, and the noise and spurious high-order wave number component cannot be suppressed sufficiently. Therefore, the reconstructed result is seriously contaminated. And, the surface velocity reconstructed from the pressure which is extrapolated with this Tikhonov filter is seriously deformed. Despite applying the data extrapolation procedure, the reconstruction error is very large because the estimated regularization parameter is too small.

On the other hand, the regularization parameter estimated from the restricted pressure by the proposed method and that estimated from the full-aperture pressure by the modified GCV method show good agreement. Therefore, the reconstruction error becomes small by the proposed method.

As a result, the availability of the proposed method is confirmed. This method may thus be recommended to promote a more quantitative analysis based on BEM-based NAH for actual large-scale structure.

ACKNOWLEDGMENTS

The author would like to thank Professor Shigeru Yoshikawa of Kyushu Institute of Design for valuable discussions and comments on this paper. The author also would like to thank Dr. Earl G. Williams at Naval Research Laboratory for helpful suggestions.

[1] J. D. Maynard, E. G. Williams, and Y. Lee, "Near-field acoustic holography I. Theory of generalized holography and the development of NAH," J. Acoust. Soc. Am. **78**, 1395–1413 (1985).
[2] E. G. Williams, H. D. Dardy, and K. B. Washburn, "Generalized near-field acoustical holography for cylindrical geometry: Theory and experiment," J. Acoust. Soc. Am. **81**, 389–406 (1987).
[3] S. Yoshikawa, E. G. Williams, and K. B. Washburn, "Vibration of two concentric submerged cylindrical shells coupled by the entrained fluid," J. Acoust. Soc. Am. **95**, 3273–3286 (1994).
[4] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, London, 1999).
[5] W. A. Veronesi and J. D. Maynard, "Digital holographic reconstruction of

sources with arbitrary shaped surfaces,'' J. Acoust. Soc. Am. **85**, 588–598 (1989).
[6] G. Borgitti, A. Sarkissian, E. G. Williams, and L. Schuets, ''Generalized near-field acoustic holography for axisymmetric geometries,'' J. Acoust. Soc. Am. **88**, 199–209 (1990).
[7] G.-T. Kim and B.-H. Lee, ''3D sound source reconstruction and field reprediction using the Helmholtz integral equation,'' J. Sound Vib. **136**, 245–261 (1990).
[8] A. Sarkissian, ''Near-field acoustic holography for an axisymmetric geometry: A new formulation,'' J. Acoust. Soc. Am. **88**, 961–966 (1990).
[9] M. R. Bai, ''Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries,'' J. Acoust. Soc. Am. **92**, 533–549 (1992).
[10] B.-K. Kim and J.-G. Ih, ''On the reconstruction of the vibro-acoustic field over the surface enclosing an interior space using the boundary element method,'' J. Acoust. Soc. Am. **100**, 3003–3016 (1996).
[11] S.-C. Kang and J.-G. Ih, ''The use of partially measured source data in near-field acoustic holography based on the BEM,'' J. Acoust. Soc. Am. **107**, 2472–2479 (2000).
[12] B.-K. Kim and J.-G. Ih, ''Design of an optimal wave-vector filter for enhancing the resolution of reconstructed source field by near-field acoustical holography (NAH),'' J. Acoust. Soc. Am. **107**, 3289–3297 (2000).
[13] E. G. Williams, B. H. Houston, P. C. Herdic, S. T. Raveendra, and B. Gardner, ''Interior NAH in flight,'' J. Acoust. Soc. Am. **108**, 1451–1463 (2000).
[14] S.-C. Kang and J.-H. Ih, ''Use of nonsingular boundary integral formulation for reducing errors due to near-field measurement in the boundary element method based near-field acoustic holography,'' J. Acoust. Soc. Am. **109**, 1320–1328 (2001).
[15] K. Saijyou and S. Yoshikawa, ''Reduction methods of the reconstruction error for large-scale implementation of near-field acoustical holography,'' J. Acoust. Soc. Am. **110**, 2007–2023 (2001).
[16] E. G. Williams, ''Continuation of acoustic nearfields,'' J. Acoust. Soc. Am. **113**, 1273–1281 (2003).
[17] P. C. Hansen, *Rank-Deficient and Discrete III-Posed Problems* (SIAM, Philadelphia, 1998).
[18] P. A. Nelson and S. H. Yoon, ''Estimation of acoustic source strength by inverse methods. I. Conditioning of the inverse problem,'' J. Sound Vib. **233**, 643–668 (2000).
[19] S. H. Yoon and P. A. Nelson, ''Estimation of acoustic source strength by inverse methods. II. Experimental investigation of methods for choosing regularization parameters,'' J. Sound Vib. **233**, 669–705 (2000).
[20] F. Augusztionvicz, ''Reconstruction of source strength distribution by inversing the boundary element method,'' in *Boundary Elements in Acoustics, Advances and Applications*, edited by O. von Estoff (WIT, Southampton, UK, 2000).
[21] P. A. Nelson, ''Some inverse problems in acoustics,'' in Proceedings of the 6th International Congress on Sound and Vibration, Copenhagen, Denmark, July 1999, pp. 7–32.
[22] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problem* (Springer, New York, 1996).
[23] A. P. Schuhmacher and P. C. Hansen, ''Sound source reconstruction using inverse BEM,'' Proceedings of Inter-Noise 2001, 2109–2114 (2001).
[24] E. G. Williams, ''Regularization method for near-field acoustic holography,'' J. Acoust. Soc. Am. **110**, 1976–1988 (2001).
[25] R. W. Gerchberg, ''Super-resolution through error energy reduction,'' Opt. Acta **21**, 709–720 (1974).

Fluid volume displacement at the oval and round windows with air and bone conduction stimulation[a)]

Stefan Stenfelt,[b)] Naohito Hato,[c)] and Richard L. Goode
Division of Otolaryngology—Head and Neck Surgery, Stanford University Medical Center, Stanford, California

(Received 23 December 2002; revised 3 November 2003; accepted 17 November 2003)

The fluids in the cochlea are normally considered incompressible, and the fluid volume displacement of the oval window (OW) and the round window (RW) should be equal and of opposite phase. However, other channels, such as the cochlear and vestibular aqueducts, may affect the fluid flow. To test if the OW and RW fluid flows are equal and of opposite phase, the volume displacement was assessed by multiple point measurement at the windows with a laser Doppler vibrometer. This was done during air conduction (AC) stimulation in seven fresh human temporal bones, and with bone conduction (BC) stimulation in eight temporal bones and one human cadaver head. With AC stimulation, the average volume displacement of the two windows is within 3 dB, and the phase difference is close to 180° for the frequency range 0.1 to 10 kHz. With BC stimulation, the average volume displacement difference between the two windows is greater: below 2 kHz, the volume displacement at the RW is 5 to 15 dB greater than at the OW and above 2 kHz more fluid is displaced at the OW. With BC stimulation, lesions at the OW caused only minor changes of the fluid flow at the RW.
[DOI: 10.1121/1.1639903]

PACS numbers: 43.64.Bt [WPF] Pages: 797–812

I. INTRODUCTION

When an air-conducted (AC) sound enters the ear, it is transmitted through the external ear canal and converted to vibrations by the tympanic membrane (TM). These vibrations are, via the middle ear ossicles, transmitted to the stapes footplate where they are converted to a sound pressure in the cochlear fluid. The general assumption about the cochlea and its contents is that the cochlear fluids can be seen as incompressible and the bone surrounding the cochlea as infinitely stiff, while the only fluid inlets and outlets of the cochlea are the round window (RW) and oval window (OW). As a consequence, the amount of fluid displaced at the OW must equal that at the RW, but of the opposite phase. Such results were reported by Kringlebotn (1995) for cadaver ears, extracted from pigs, when stimulated by AC. In the same investigation, Kringlebotn reported a 5-dB difference between the OW and RW fluid displacement for the same measurements conducted in a human temporal bone specimen.

The transmission of a bone-conducted (BC) stimulus to the cochlea is different. The important pathways for BC stimulation are believed to be (1) sound radiated into the external ear canal and middle ear cavity, which is transmitted to the cochlea via the TM and the middle ear ossicles; (2) inertial effects of the middle ear ossicles; (3) inertial effects of the fluid in the cochlea; and (4) compression and expansion of the petrous bone encapsulating the cochlea. It is clear that, if the volume of the whole cochlea can be changed by BC stimulation, the fluid displacement at the RW and OW can differ. According to the theory, this can occur only at frequencies above 0.8–1.0 kHz (Håkansson *et al.*, 1994). Below these frequencies, the skull moves as a rigid body and there is no expansion and compression of the cochlea.

It has recently been shown that a sound pressure introduced in the cerebrospinal fluid (CSF) is transmitted to the cochlea and gives rise to a hearing sensation without causing enough vibration of the skull bones to be considered normal BC transmission (Freeman *et al.*, 2000; Sohmer *et al.*, 2000). This indicates that there are fluid pathways from the CSF to the cochlea which are patent to sound pressure transmission. It has previously been shown that a static pressure can be transmitted from the CSF to the cochlear fluid (Carlborg *et al.*, 1982; Yoshida and Uemura, 1991). This pressure transmission was attributed to the cochlear and vestibular aqueducts. In the human, the cochlear aqueduct is an approximately 10-mm-long, narrow canal. It is about 0.1 mm in diameter at its narrowest part, and is often reported to contain tissue. However, the patency of the cochlear aqueduct in the human has been disputed (Ritter and Lawrenece, 1965; Palva and Dampert, 1969; Rask-Andersen *et al.*, 1977; Nomura, 1984; Gopen *et al.*, 1997).

That the aqueducts could transmit fluid in and out of the cochlea was early proposed by Békésy (1936). Tonndorf (1962), when experimentally immobilizing the RW, found only a reduction in the AC sensitivity of 20 to 30 dB in cats. He attributed the residue hearing to incomplete immobilazation of the RW. However, in the same study he reported that no significant change was found in either BC or AC sensitivity when the RW was obstructed, as long as there is an open cochlear aqueduct. Groen and Hoogland (1958), when

[a)]Portions of this work were presented at the 141st meeting of the Acoustical Society of America, Chicago, Ill, 2001.

[b)]Present address: Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Göteborg, Sweden. Electronic mail: stenfelt@s2.chalmers.se

[c)]Present address: Department of Otolaryngology, Ehime University School of Medicine, Shitsukawa, Shigenobu-cho, Onsen-gun, Ehime, Japan.

investigating a patient with otosclerosis of the RW, found that hearing by BC deteriorates very little in the low frequency region, provided the cochlear aqueduct and the inferior cochlear vein remain open. At higher frequencies, they found thresholds to deteriorate by 6 dB/oct, which they attributed to the increased impedance with frequency of the fluid canal. Carhart (1962) reported that the typical BC threshold loss with otosclerosis of the stapes is 15–25 dB at around 2 kHz, and 5–10 dB at 1 and 4 kHz. At higher and lower frequencies, no effect on the BC hearing is seen. Everberg (1968) found close to normal BC thresholds in patients with no oval window due to malformations. This indicates that the fluid flow at the two windows can differ when stimulation is by BC.

Since it seems to be somewhat unclear whether there is a contribution from inlets and outlets in the cochlea, besides those of the OW and RW, to the fluid dynamics of the cochlea, the following experiment was devised. The fluid volume displacement at the OW (stapes footplate) was measured and compared with RW fluid displacement in the same temporal bone specimen. This was done for both AC and BC stimulation. Further, with BC stimulation, the stapes footplate was glued to the promontory bone to simulate an otosclerotic condition of the stapes footplate, and to reveal its influence on the fluid volume displacement at the RW. A hole was then drilled in the stapes footplate, which removed much of the impedance at the OW, thus revealing some of the fluid dynamics in the cochlea.

II. MATERIALS AND METHODS

A. Temporal bone experiments

For this investigation, 15 human temporal bone specimens were studied (12 male and 3 female), with an average age of 72 and a range of 56 to 93 years. The temporal bones were extracted from human cadavers, within 48 h of death, using a Schuknecht bone saw at the time of autopsy. The temporal bone specimens were wrapped in gauze, placed in a 1:10000 merthiolate solution in normal saline, and stored at 5 °C. All measurements on individual bones were conducted on the same day, within 6 days of death. The TM and middle ear in each bone were inspected using an operating microscope; bones with abnormal TMs or middle ears were excluded from the study.

The temporal bones for AC and BC testing were prepared differently. Temporal bones investigated with AC stimulation were prepared as follows. The attached connective tissue was removed, and the bony wall of the external ear canal was drilled down to 2 mm from the tympanic annulus. A simple mastoidectomy (opening of the mastoid antrum and removal of the mastoid air cells) and a posterior hypotympanotomy (widely opening the facial recess) were performed, including removal of the mastoid portion of the facial nerve and surrounding bone, to provide a good view of the stapes and stapes footplate. To obtain a good view of the RW membrane, the RW niche had to be slightly widened. The ossicular ligaments, chorda tympani, tensor tympani, and stapedius muscle were left intact. A 25-mm-long plastic artificial external ear canal of 8.5-mm i.d. was placed against the bony ear canal remnant; it was placed so that the axis of the ear canal was approximately perpendicular to the TM. The temporal bone was then sealed in latex and covered with dental cement [Fig. 1(a)]. This prevented the bone from drying during the time needed for the measurement.

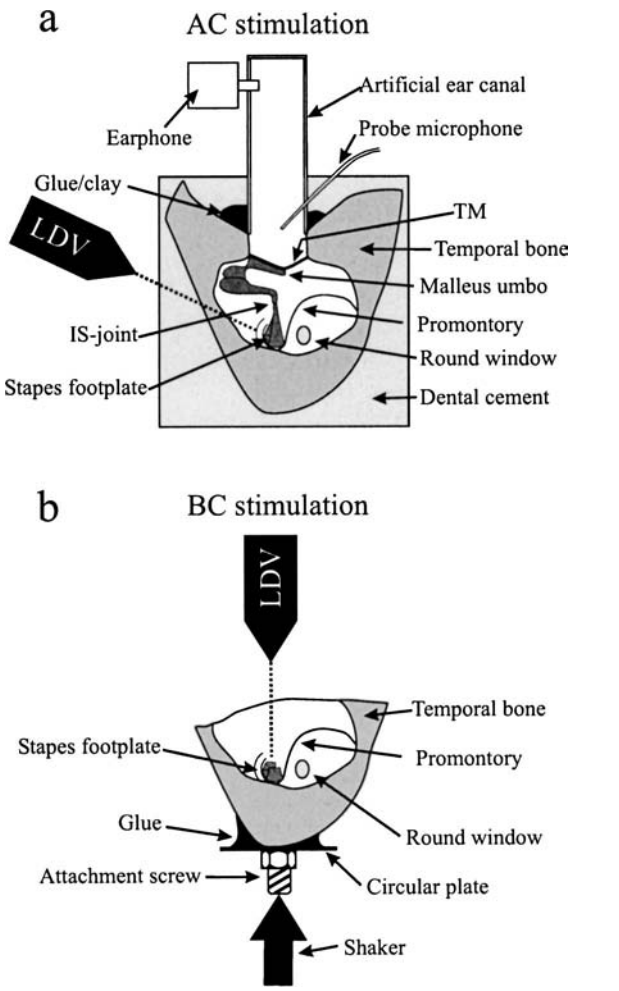


FIG. 1. (a) Temporal bone specimen preparation for measurements with AC stimulation. A 25-mm-long artificial ear canal is glued to the specimen. The earphone at the lateral end of the ear canal provides sound stimulus that is measured by the probe microphone approximately 2 mm in front of the TM. Vibrations of the stapes footplate and RW membrane is measured by the LDV through the facial recess. The whole specimen is embedded in dental cement. (b) Temporal bone specimen preparation for measurements with BC stimulation. The bone around the ear canal together with the TM, malleus, incus, and stapes superstructure are removed to give a good view of the stapes footplate, cochlear promontory, and RW. A circular plate with an attachment screw is glued to the specimen so the main vibration direction is the same as the normal in and out vibration direction of the stapes footplate. The vibrations of the stapes footplate are measured by the LDV perpendicular to the stapes footplate and the vibrations of the RW membrane are measured by the LDV perpendicular to the RW membrane (not shown in the figure). Part of the bone in the RW niche is removed in order to achieve a perpendicular view of the RW membrane.

If the temporal bone specimen was to be tested with BC stimulation, the attached connective tissue was removed, as well as the TM, malleus, incus, and the stapes superstructure [Fig. 1(b)]. Care was taken not to damage the stapes footplate or the annular ligament that connects the stapes footplate to the oval window. Then the bony wall of the external ear canal and middle ear was removed, leaving only the bone surrounding the cochlea and the semicircular canals. This

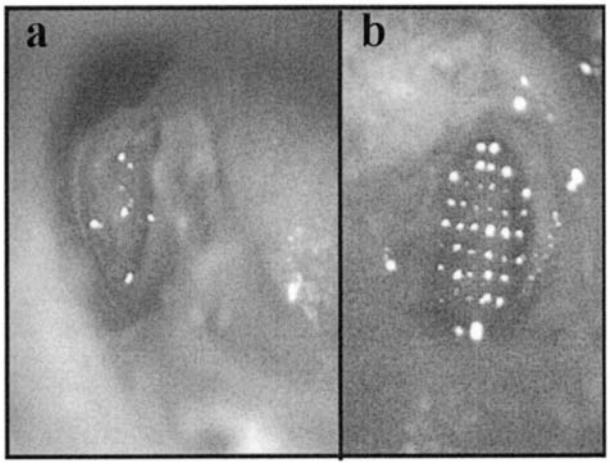


FIG. 2. (a) Photograph of the stapes footplate and the five reflective targets used for the LDV measurements when stimulation was by AC. When BC stimulation was used, the whole superstructure of the stapes was removed. (b) Photograph of the RW. A grid of reflective targets with a spacing of 0.2 mm was used to measure the RW membrane vibrations. Four reflective targets were placed on the bone just outside the RW. These were used to determine the vibration of the bone surrounding the RW when stimulation was by BC. A similar approach was also used to measure the vibration of the bone surrounding the oval window when stimulation was by BC.

enabled a perpendicular view of the stapes footplate and the RW membrane for the measurements with the laser Doppler vibrometer (LDV). In addition, it removed excessive mass of the temporal bone to ensure translational motion of the specimen during the measurements (Stenfelt *et al.*, 2002). A flat circular aluminum plate, with a diameter of 15 mm and a screw attachment, was glued to the prepared temporal bone with cyanoacrylate glue (Garf Reef glue™, Garf Inc, Boise, ID). This plate was later screwed to the B&K type 4810 mini-shaker (Brüel and Kjær, Nærum, Denmark), ensuring a rigid connection to the mini-shaker and a translational vibration of the specimen. The plate was glued to the temporal bone specimen along the axis of the normal in-and-out motion of the stapes footplate [Fig. 1(b)].

To achieve good reflection of the laser beam, small reflective glass micro-spheres ($\varnothing$ 5 μm) were positioned at the measurement points. The positioning of these micro-spheres was similar for the AC and BC measurements. Five micro-spheres were positioned on the stapes footplate for the measurement of the volume displacement at the oval window: one in the center and one at the posterior, superior, anterior, and inferior side of the footplate. A matrix of micro-spheres, with circa 0.2-mm spacing, was placed on the RW membrane (Fig. 2). Depending on the size of the RW, somewhere between 27 and 40 equally spaced micro-spheres were placed on the RW membrane. For both the OW and RW measurement, four additional micro-spheres were approximately equally spaced, placed on the circumference of the RW and OW, on the promontory bone [Fig. 2(b)]. The velocity of these four positions was used to calculate the bone velocity in the plane of either RW or OW. These two bone velocities were used to calculate the velocity difference between the RW membrane and the promontory bone as well as the stapes footplate and the promontory bone.

Afterwards, a calibrated photograph was taken of the stapes footplate and the RW membrane to show clearly the micro-sphere positions and the edges of the stapes footplate and the RW membrane. The coordinates of the micro-spheres were stored to be used to calculate the motion of the stapes footplate and the RW membrane. In addition, the photographs were used to calculate the areas of the stapes footplate and the RW membrane. For the 15 temporal bones used in this study, the average area of the stapes footplate was 3.85 mm^2 (3.00–5.30 mm^2) and the average area of the RW was 2.39 mm^2 (1.68–3.67 mm^2).

B. Whole head measurement

To verify that the results obtained in the temporal bone specimens are not due to artifacts from extracting the temporal bones from the heads, one cadaver human head was used. It is here assumed that extracting temporal bones influences the response more with BC stimulation than with AC stimulation; BC stimulation involves more complex phenomena that can be altered by the extraction than AC stimulation does. Therefore, the measurement of RW and OW fluid volume displacement in a whole head is only done with BC stimulation.

No history of the head was known, except that it was a male and between 60 and 70 years old. It was severed between the third and fourth vertebra at the time of autopsy. Its weight was 3.62 kg and had a circumference of 56 cm measured across the middle of the forehead, just above the ear canal openings and across the occiput. Apart from being severed, the head was left intact, i.e., no cutting or drilling of the head during the autopsy, and the soft tissues and brain were also left intact.

At the time of measurement, the ear was prepared similar as the temporal bone specimens stimulated by BC. The RW and OW volume displacements in the cadaver head were measured and calculated in a way similar to that described for the temporal bone specimen with BC stimulation. For BC stimulation in the cadaver head, a small transducer was rigidly attached by means of a screw connection to the head at the mastoid bone (Stenfelt *et al.*, 2000).

C. Measurement system

The stimulation, either as sound pressure in the artificial ear canal or vibration of the temporal bone specimen, was provided by PC based software, SYSid 6.5, using a DSP-16+ signal processing card (www.sysid-labs.com). The output from the computer was fed through a power amplifier (D-75, Crown, Elkhart, IN) and, for AC stimulation, to a receiver in the artificial ear canal (Tibbets 83-13A/024, Tibbets Industries, Camden, ME). For BC stimulation, the output from the power amplifier was fed to a B&K type 4810 mini-shaker. The mini-shaker was rigidly coupled to the temporal bone by a threaded connector glued to the temporal bone [Fig. 1(b)]. The vibrations, either of the stapes footplate, the RW membrane, or the promontory bone, were measured with an LDV, the HLV-1000 (Polytec, Waldbronn, Germany). Its sensor head was mounted with a joystick-controlled mirror on an operating microscope, enabling easy control of the laser beam. A probe tube microphone (ER-7C,

Etymotic Research, Elk Grove Village, IL), with its opening positioned 2 mm from the TM in the artificial ear canal, measured the ear canal sound pressure.

First, with AC stimulation, the normal middle ear transfer function of the temporal bone (malleus umbo to stapes footplate) was measured to verify the specimen to be within normal limits. The five positions on the stapes footplate were then measured, and, last, the positions on the RW membrane. The total time for the measurements was about one-half hour. The results of the measurements were stored for later computation of the volume displacement at each window. The angle between the laser beam and the perpendicular view, during the measurement of the stapes footplate vibration, was 57°; it was 30° when measuring the RW membrane motion. These values were taken into account when calculating the volume displacement: the results were converted to a motion perpendicular to the stapes footplate and the RW membrane.

The procedure was similar when the stimulation was by BC. However, with BC stimulation, the laser beam was aimed perpendicular to the surface measured, either the stapes footplate or the RW membrane. This minimized the influence of motion from directions other than that of the stimulation. Thus, no correction for misalignment had to be made for the measurements with BC stimulation. The promontory bone vibration close to the OW and RW was also measured with the LDV. These data were used when calculating the differential motion between the stapes footplate, or RW membrane, and the promontory bone. For BC stimulation, the vibration of the stapes footplate and the RW were measured for three conditions of the cochlea. First, measurements of the footplate and RW membrane motion were made for a normal intact cochlea. Next, the same measurements were repeated after the stapes footplate was glued into the RW niche with cyanoacrylate glue (Garf Reef Glue™). Then, a hole of approximately ∅ 1 mm was made in the stapes footplate to simulate a fenestration operation, and the RW membrane vibration was remeasured. In total, the three types of measurements took about 1 1/2 h to complete. To reduce drying effects, the specimens were sprayed with saline every 10 min.

D. Calibration

To limit high frequency noise or overloading by low frequency vibrations, HLV-1000 filters were used. The low-pass filter cutoff frequency was set at 15 kHz and the high-pass cutoff frequency at 100 Hz. These settings, which affected the frequency response of the HLV-1000, were calibrated with a B&K type 4371 accelerometer. Below 10 kHz, this accelerometer has, according to the manufacturer, a maximum level deviation of 0.2 dB and a maximum phase deviation of 5°. The accelerometer was mounted on the B&K 4810 mini-shaker, and the laser was aimed at the surface of the accelerometer in a line perpendicular to the accelerometer surface. This was used for calibration of the HLV-1000 for the frequency range 0.1 to 10 kHz (50 frequencies/decade). The ER-7C probe tube microphone was calibrated for a B&K type 4138 $\frac{1}{8}$-in. microphone. The sensitivity of the B&K microphone was first determined in a B&K type 4230 sound level calibrator. Both microphones were then placed 1 mm apart in a small cavity, a sound was introduced, and the calibration curve of the ER-7C obtained for the frequency range 0.1 to 10 kHz (50 frequencies/decade).

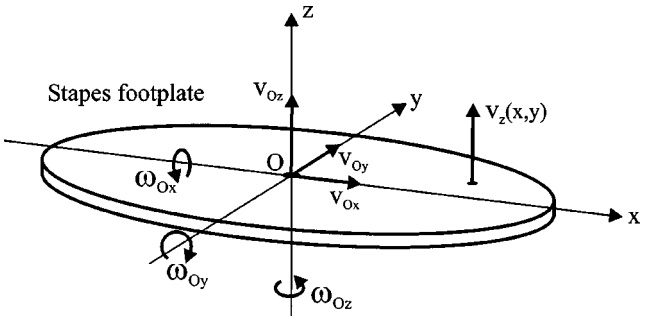


FIG. 3. A model of the stapes footplate motion is shown in a relative coordinate system where the origin O is at the center of the footplate. Shown are the translational velocity components at the center (v_{Ox},v_{Oy},v_{Oz}) and the angular velocity components (ω_{Ox},ω_{Oy},ω_{Oz}). Due to the constraints of the annular ligament, the velocity components in the plane of the footplate can be excluded and the resulting velocity at a position $v_z(x,y)$ is a function of the translational velocity v_{Oz} and the two rotational components ω_{Ox} and ω_{Oy}.

E. Calculation of volume displacement

1. AC stimulation

The calculation of the stapes footplate motion from the five-point measurement was described in Hato *et al.* (2003). Here, the same equations are used to extract the fluid displacement at the OW. The motion at any point *p* on a rigid body can be described as

$$\mathbf{v_p} = \mathbf{v_O} + \boldsymbol{\omega_O} \times \mathbf{r_{Op}}, \tag{1}$$

where $\mathbf{v_p}$ is the three component velocity (v_{px},v_{py},v_{pz}) at position $\mathbf{p}$, $\mathbf{v_O}$ is the three component translational velocity (v_{Ox},v_{Oy},v_{Oz}) at position $\mathbf{O}$ (origin), $\boldsymbol{\omega_O}$ is the three component angular velocity (ω_{Ox},ω_{Oy},ω_{Oz}) about position $\mathbf{O}$, and $\mathbf{r_{Op}}$ is the vector between position $\mathbf{O}$ and $\mathbf{p}$. The stapes footplate is assumed to move as a rigid body. The stapes footplate is positioned in the OW fenestra by the annular ligament. Since this ligament constrains the velocity in the plane of the footplate (v_{Ox},v_{Oy}) there is less motion in the plane of the footplate than perpendicular to the footplate surface (v_{Oz}). Similarly, the rotation ω_{Oz} perpendicular to the plane of the footplate is considerably lower than the rotations in the plane of the footplate (ω_{Ox} and ω_{Oy}). Equation (1) can be simplified to a one-direction velocity (Fig. 3):

$$v_{pz}(x,y) = v_{Oz} + \omega_{Ox} y - \omega_{Oy} x, \tag{2}$$

where *x* and *y* can be regarded as constants due to the small vibrations. Moreover, it is only v_{pz} that contributes to the fluid flow at the OW. Combining different positions *p*, Eq. (2) can be written in matrix form as

$$\begin{bmatrix} v_{1z} \\ v_{2z} \\ \vdots \\ v_{nz} \end{bmatrix} = \begin{bmatrix} 1 & y_1 & -x_1 \\ 1 & y_2 & -x_2 \\ \vdots & \vdots & \vdots \\ 1 & y_n & -x_n \end{bmatrix} \begin{bmatrix} v_{Oz} \\ \omega_{Ox} \\ \omega_{Oy} \end{bmatrix} \tag{3}$$

or

$$\mathbf{V}_z = \mathbf{BA}. \tag{4}$$

Due to measurement inaccuracies the measured velocity v_{pz} can deviate from its true value, meaning that Eq. (3) is not exactly for $n>3$. But estimated (minimum square) values for the complex parameters v_{Oz}, ω_{Ox}, and ω_{Oy} can be obtained from the velocities measured at the five different target positions according to

$$\begin{bmatrix} \hat{v}_{Oz} \\ \hat{\omega}_{Ox} \\ \hat{\omega}_{Oy} \end{bmatrix} = (\mathbf{B^T B})^{-1} \mathbf{B^T V}_z, \tag{5}$$

where $^{-1}$ means matrix inversion and $^{\mathrm{T}}$ is the transpose of the matrix. These three estimated parameters can then be used to calculate the motion of the of the stapes footplate according to Eq. (2). The resulting velocity estimate at a position (x,y) on the stapes footplate is then

$$\hat{v}(x,y) = \hat{v}_{Oz} + \hat{\omega}_{Ox} y - \hat{\omega}_{Oy} x, \tag{6}$$

In order to obtain the fluid volume displacement at the OW, the estimated velocity of the footplate $\hat{v}(x,y)$ at a position (x,y) is converted to a displacement $\hat{d}(x,y)$ by integration with respect to time:

$$\hat{d}(x,y) = \int \hat{v}(x,y)dt. \tag{7}$$

The volume displacement produced by the whole stapes footplate was estimated by dividing the footplate into N segments of 0.1×0.1 mm. For each of these segments, the displacement at the center of the segment is estimated by Eqs. (6) and (7). The volume displacement for each segment is calculated as the displacement at its center times $10^{-8}\,\mathrm{m}^2$. The volume displacement of the fluid at the OW is the complex summation of the fluid volume displacement of all footplate segments:

$$d_{OW} = \sum_{n=1}^{N} d_{n,footplate} \cdot 10^{-8}, \tag{8}$$

where d_{OW} is the volume displacement at the OW and $d_{n,footplate}$ is the volume displacement of the nth segment of the stapes footplate.

The volume displacement of the fluid at the RW was estimated in a more direct manner. The area of the RW was divided into M segments by the matrix of the micro-spheres placed on the RW membrane; the size of each was approximately 0.2×0.2 mm. The area of each segment ($A_{n,RW}$) was measured from the photograph taken of the RW membrane. The volume displacement of each segment was its area times the displacement of that segment. The displacement of each segment was determined as the velocity of the target in the center of the segment, measured by the LDV, and then integrated with respect to time to give displacement. The displacement of each segment was a complex number containing the amplitude and phase of its displacement; the total fluid displacement at the RW (d_{RW}) was obtained as the complex sum of all displacements as

$$d_{RW} = \sum_{n=1}^{M} d_{n,RW} \cdot A_{n,RW}, \tag{9}$$

where $d_{n,RW}$ is the displacement of segment n, $A_{n,RW}$ is the area of segment n taken from the photograph, and M is the total number of segments. Some of the micro-spheres were placed just outside the edge of the RW. Since the calculated displacements of these were close to zero, they were excluded from the summation of the RW displacement.

2. BC stimulation

The calculation of volume displacement at the OW and RW was similar for AC and BC measurements. However, with BC stimulation the velocity of the bone surrounding the OW and RW was measured as well. The velocities in Eqs. (1)–(6) of the BC measurements were calculated as the difference between the RW membrane, or stapes footplate, and the promontory bone:

$$v_{BC,FP} = v_{FP} - v_{Promontory,FP}, \tag{10}$$

$$v_{BC,RW} = v_{RW} - v_{Promontory,RW}. \tag{11}$$

For the calculation of the relative footplate velocity [Eq. (10)], the promontory bone velocity ($v_{Promontory,FP}$) close to the OW was measured in a line perpendicular to the stapes footplate. Similarly, for the calculation of the relative RW membrane velocity [Eq. (11)], the promontory bone velocity ($v_{Promontory,RW}$) close to the RW was measured in a line perpendicular to the RW membrane. This means that the beam of the LDV had the same orientation for the measurement of the stapes footplate as it had for the promontory bone measurement adjacent to the OW. Similarly, the LDV beam had the same orientation for the RW membrane measurement as for measurements of the promontory bone adjacent to the RW. The four positions on the promontory bone adjacent to either the OW or RW were used to calculate the bone velocity for the stapes footplate and RW membrane measurements. To obtain this bone velocity from the four measurement positions on the promontory bone, the same scheme as in Eqs. (1)–(6) was used; however, only four measurement positions were used. This means that the relative velocity according to Eqs. (10) and (11) were calculated for each position as

$$v_{BC,FP}(x,y) = v_{FP}(x,y) - v_{Promontory,FP}(x,y), \tag{12}$$

$$v_{BC,RW}(x,y) = v_{RW}(x,y) - v_{Promontory,RW}(x,y), \tag{13}$$

where (x,y) describes a coordinate system for the OW and RW, respectively.

After these relative velocities were calculated, the volume displacements at the OW and the RW with BC stimulation were obtained in a way similar to that for AC stimulation just described. The difference was that, for BC stimulation, the promontory bone velocity in the direction of stimulation was used as input reference, while the sound pressure at the TM was used as the input reference for the AC measurements.

F. Error analysis

One source of error in these measurements is due to uncertainties determining the position of the targets. For the volume displacement computation, the center of each target was determined in a calibrated photograph. The accuracy of the target position is estimated to be within 5 μm. That uncertainty does not cause large errors provided the spatial distribution of the motion is smooth; this is true for the stapes footplate motion (stiff plate) and the low-frequency motion of the RW membrane. At the high frequencies, the RW membrane vibration breaks up in several modes and the motion is complex. Thus, the uncertainty in target position can cause errors in the volume displacement estimation of the RW at the high frequencies (above 8 kHz). However, the targets are close enough on the RW membrane to fulfill the Nyquist two-dimensional sampling theorem for frequencies up to 10 kHz.

The anatomy of the RW membrane is another source for bias when estimating the fluid displacement. The RW membrane is not flat; it has a slightly convex curvature. This curvature was pronounced in some ears while others had a nearly flat RW membrane. The greatest misalignment between a part of the membrane and the line of measurement was 30°, which corresponds to a bias of about 1 dB. Approximately one-third of the RW membrane could be misaligned; i.e., the error in the estimation of the volume displacement due to the curvature of the RW membrane would be less than 0.5 dB.

When AC stimulation was used, the angle between the LDV beam and the direction of the in-and-out motion of the stapes footplate (denoted *z* direction previously) is 57°. This angle was determined using a mass in a thin string attached to the microscope to determine the vertical line. Then a protractor was used to set the angle between the LDV beam and the vertical line to 57°. The stapes footplate was aligned in the horizontal plane using a temporal bone holder. The alignment was not verified and it was estimated to be within ±5°. This means that the uncertainty due to error of the alignment when measuring the stapes footplate vibration with AC stimulation is +1.1 to −1.3 dB. The RW membrane was in a similar way measured at an angle of 30°. A 5° uncertainty in the angle gives an uncertainty of +0.4 to −0.5 dB for the RW membrane measurements.

The length of the LDV beam is approximately 300 mm. No correction in the angle was made for measurements at different positions on the stapes footplate or on the RW membrane. The largest distance between any two targets on either the stapes footplate or RW membrane is less than 3 mm, which corresponds to a deviation in the angle of less than 0.5°. The uncertainty in the measurement due to this is therefore considered insignificant.

When BC stimulation was used, the LDV beam was aimed perpendicular to the stapes footplate surface or the RW membrane surface. With the temporal bones, a misalignment of ±5° means an uncertainty of less than 0.1 dB. When the whole head measurement was conducted no holder was used but the head was placed on a pillow. Therefore, the alignment was slightly worse and estimated to be within ±10° which means that the error due to misalignment for the whole head is bound by 0.2 dB.

The greatest uncertainty of the BC data comes from the calculation of $v_{BC,FP}$ and $v_{BC,RW}$. This uncertainty can be calculated from the errors in the measurement of v_{FP}, $v_{Promontory,FP}$, v_{RW}, and $v_{Promontory,RW}$. If Eq. (12) is rewritten as

$$\frac{v_{BC,FP}}{v_{Promontory,FP}} = \left(\frac{v_{FP}}{v_{Promontory,FP}} - 1 \right),$$

the uncertainty of $v_{BC,FP}$ can be related to the velocity relation between $v_{BC,FP}$ and the bone. The test-retest variability of the measured velocities was less than 0.2 dB. This gives an uncertainty for the quotient $v_{FP}/v_{Promontory,FP}$ of 0.4 dB or 4.7%. Therefore, due to the subtraction in the above equation, the uncertainty results in a limit of −26 dB or a ratio of 0.047. This is a conservative limit since the −26-dB level is a worst case scenario; reliable data is typically obtained for levels well below this limit. As reported in Stenfelt *et al.* (2002), no increase in the spread of the data was found for levels down to −40 dB using the same measurement technique as in this study, but it had even worse test-retest variability than in this study.

For the BC measurements in this study, a limit at −26 dB corresponds approximately to a lower frequency limit of 1.0 kHz, and a limit of −40 dB corresponds approximately to a lower frequency limit of 0.5 kHz (Sec. III B, Fig. 8). However, due to the stiffness of the RW membrane and the annular ligament in the OW, the low-frequency behavior of $v_{BC,FP}/v_{Promontory,FP}$ and $v_{BC,RW}/v_{Promontory,RW}$ is rising at 40 dB per decade. This is apparent for frequencies above 0.3 kHz (Fig. 8) and the BC result in this study is considered valid for frequencies above 0.3 kHz.

III. RESULTS

A. Stimulation by AC

Although a detailed analysis of the vibration modes of the stapes footplate and the RW membrane with AC stimulation is not given here, a brief description is presented. A thorough description of the stapes footplate motion is presented in Hato *et al.* (2003); the results were similar to this study. The stapes footplate moved predominantly with a pistonlike motion at frequencies below 2 kHz [Figs. 4(a) and (b)]. At higher frequencies, additional rotational motion of the stapes footplate rose to a maximum around 6 kHz; above this frequency, the stapes footplate showed complex vibration pattern. For frequencies below 1.5 kHz, the RW membrane showed a single mode of vibration [Figs. 4(c) and (d)]; above 1.5 kHz the RW membrane vibration breaks up into several modes.

With AC stimulation, the average fluid volume displacement at the OW and RW for seven temporal bone specimens, when the input is a sound pressure of 80 dB SPL at the TM, is shown in Fig. 5. At low frequencies, the volume displacement is close to 3×10^4 μm^3 at both windows; the volume displacement at the RW is slightly lower than at the OW. Above 1 kHz, the volume displacement at both windows is similar; it falls off by approximately −12 dB/oct. The mag-

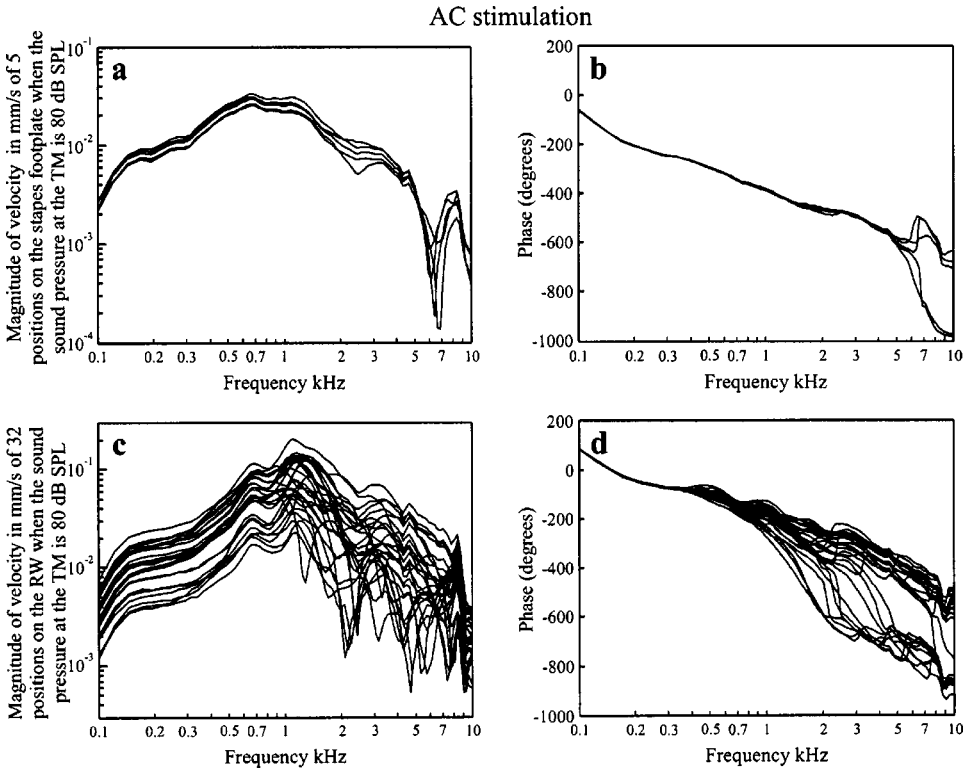


FIG. 4. The velocity magnitude (a) and phase (b) of the five positions [Fig. 2(a)] on the stapes footplate in one temporal bone specimen when a sound stimulation of 80 dB SPL is presented at the TM. With the same stimulation in the same specimen, the velocity magnitude (c) and phase (d) of 32 positions on the RW membrane are shown [Fig. 2(b)]. The velocity magnitude is shown as peak values and the phase is measured relative to the ear canal sound pressure. The spatial resolution on the RW membrane is 0.2 mm and the frequency resolution of the measurements are 50 points/decade.

nitude difference at the two windows is less than the standard deviation of the measurements (vertical bars in Fig. 5) for the whole frequency range. The phase rolls off at approximately the same rate at both windows, separated by nearly 180°.

The levels and phases of the fluid volume displacement at the RW are compared with those of the OW in Fig. 6. Both the individual data from the seven specimens and the average data are presented. The average volume displacement level at the RW is just below that at the OW. For frequencies below 1.0 kHz the difference is within 3 dB for the average and within 5 dB for the individual data; above 1.0 kHz the average result is close to 0 dB while the specimen data are bounded by +3 dB to −4 dB. The average phase of the RW volume displacement compared with the OW volume displacement is close to −180 degrees, except at frequencies above 7 kHz, where the phase difference decreases. The individual phase differences are within 30° of −180° for frequencies below 7 kHz. Above 7 kHz there is a general trend of less phase difference.

B. Stimulation by BC

With BC stimulation, the absolute velocity magnitude from one temporal bone specimen, as measured by the LDV,

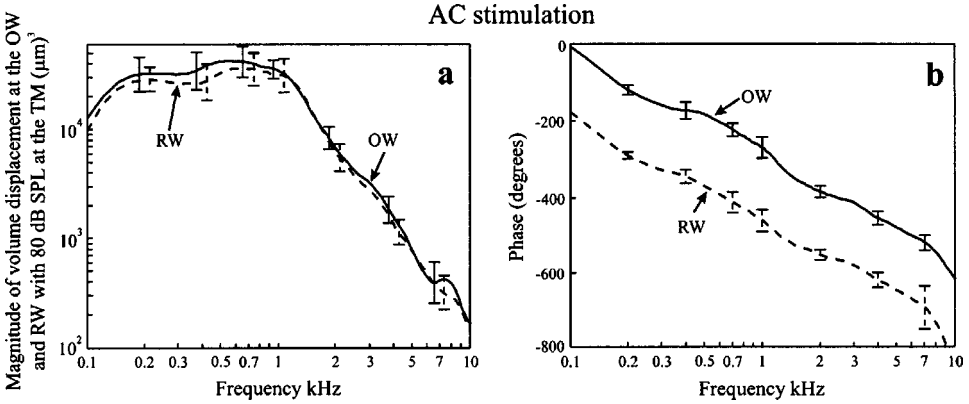


FIG. 5. Average peak magnitude (a) and phase (b) of the fluid volume displacement for seven temporal bone specimens. The specimens are stimulated in the artificial ear canal by an AC sound of 80 dB SPL at the TM. The volume displacement at the OW is given by the solid line, and the volume displacement at the RW is given by the dashed line. The phase is calculated relative to the input sound at the TM. The vertical bars give the standard deviation of the volume displacement. The frequency resolution is 50 points/decade.

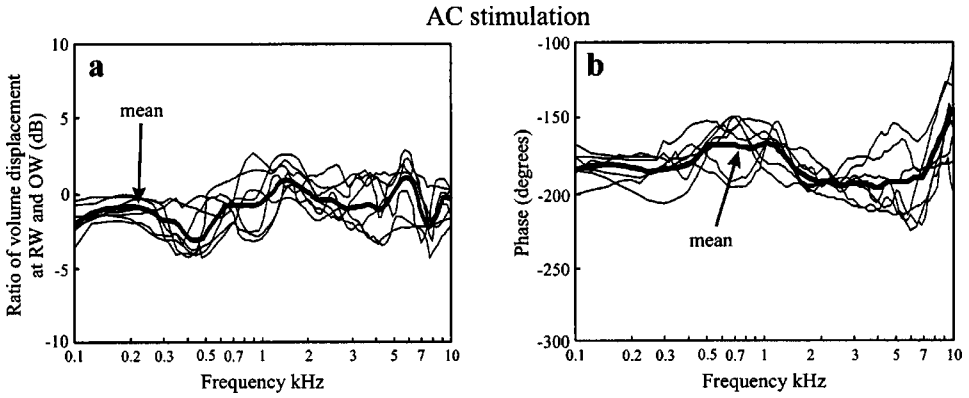


FIG. 6. Level (a) and phase (b) of the ratio of fluid volume displacement between the RW and the OW when stimulation is by AC. The ratio is calculated as the volume displacement at the RW divided by the volume displacement at the OW. All individual results for the seven temporal bone specimens (thin lines) as well as the average (thick line) are shown. For the level result, the average is calculated as the mean of the dB values. The frequency resolution is 50 points/decade.

is shown in Fig. 7 for four positions: (1) bone close to the OW, (2) center of the stapes footplate, (3) bone close to the RW, and (4) center of the RW membrane. The stimulation axis is aligned with the in-and-out motion of the stapes footplate. As a consequence, the absolute bone velocity is greater when measured at the OW than at the RW. Similar results were seen in all eight specimens measured when stimulation was by BC. Figure 7(b) gives the phase of the velocity between the center of the stapes footplate and the bone close to the OW, as well as that between the center of the RW membrane and the bone close to the RW. The figure shows that, for frequencies below 1.0 kHz, the velocity of the stapes footplate is similar in magnitude and phase to that of the surrounding bone. Due to the calculation of the difference velocity, the result at the OW is sensitive to noise in the measurement at low frequencies. The velocity difference between the RW membrane and the bone close to the RW is greater, in both magnitude and phase. It is only for frequencies below 0.3 kHz that the velocities of the membrane and bone are similar.

The relative velocity between the targets on the stapes footplate and the promontory bone in one temporal bone specimen when the stimulation of the specimen is 1 mm/s by BC are shown in Figs. 8(a) and (b). With the same stimulation in the same specimen, Figs. 8(c) and (d) shows the magnitude and phase between the targets on the RW membrane and the promontory bone. The relative velocity at the stapes footplate [Fig. 8(a)] rises with approximately 40 dB/decade above 0.3 kHz; the phase of the relative velocity at the five positions are similar [Fig. 8(b)]. The magnitude of the relative velocity of the RW membrane rises with about 40 dB/decade between 0.3 and 3.0 kHz; it flattens out above 3.0 kHz. There is more variation in the magnitude at the RW [Fig. 8(c)] than at the OW [Fig. 8(a)]. There are several peaks and valleys in the RW relative velocity magnitude above 1.0 kHz [Fig. 8(c)]. The relative velocity of the positions on the RW shows similar phase below 1.2 kHz; above 1.2 kHz the phase differs between the positions. This pattern indicates complex vibration of the RW membrane for frequencies above 1.2 kHz. Similar complex membrane vibration was found when AC stimulation was used [Figs. 4(c) and (d)].

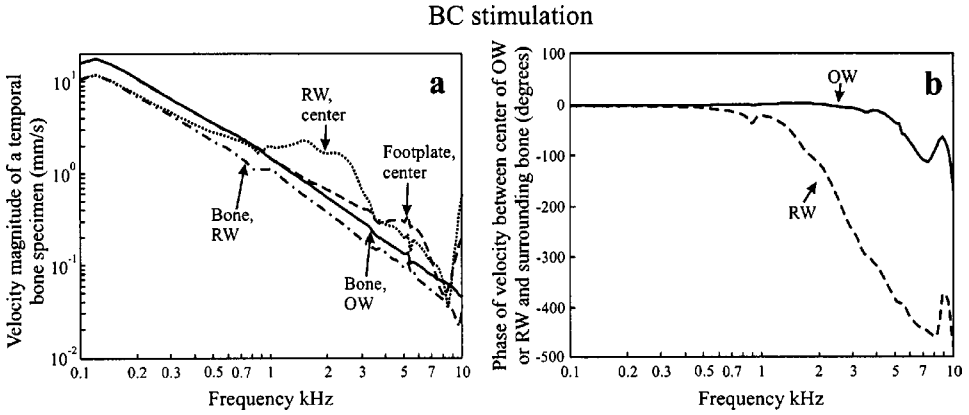


FIG. 7. (a) Magnitude of the velocity in mm/s (peak) from one temporal bone specimen when stimulation is by BC. The measurement positions are the promontory bone close to the OW (solid line); the center of the stapes footplate (dashed line); the bone next to the RW (dashed-dotted line); and the center of the RW membrane (dotted line). The OW data are measured in a line perpendicular to the stapes footplate, and the RW data are measured in a line perpendicular to the RW membrane. The stimulation axis is aligned with the normal in-and-out motion of the stapes footplate (same as the measurement direction of the OW). (b) Velocity phase of the stapes footplate center relative to the surrounding bone (solid line), and between the center of the RW membrane and the surrounding bone. All eight temporal bone specimens stimulated by BC showed similar results. The frequency resolution is 50 points/decade.

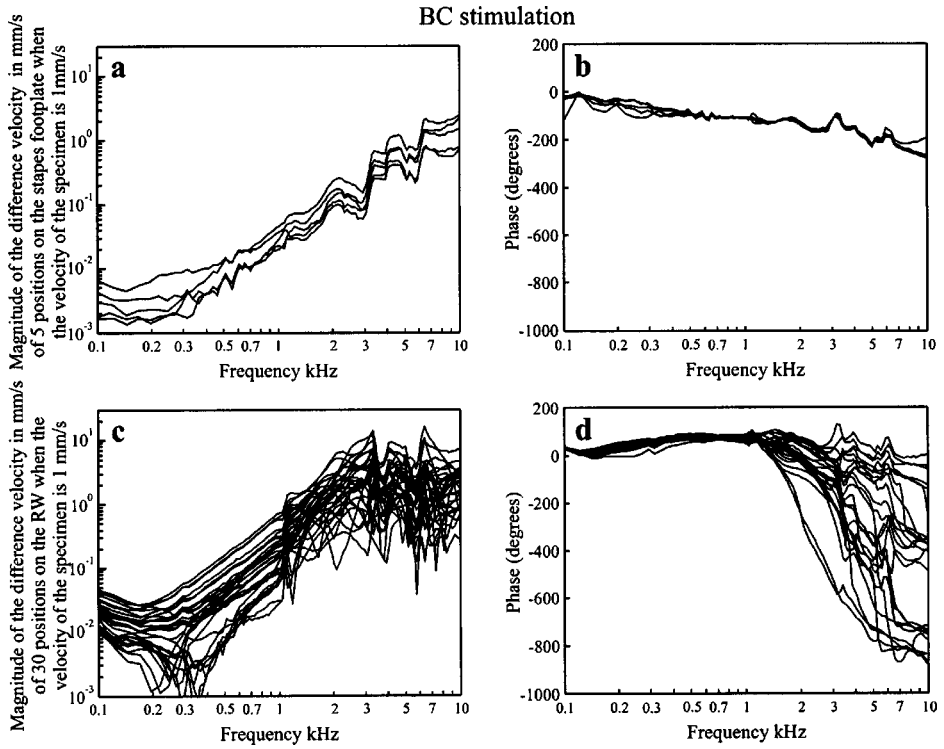


FIG. 8. From one temporal bone specimen, the difference velocity peak magnitude (a) and phase (b) of the five measurement positions [Fig. 2(a)] on the stapes footplate when the stimulation velocity of the specimen is 1 mm/s in line of the normal in-and-out motion of the stapes footplate. The difference velocity is calculated as the velocity of the stapes footplate minus the velocity of the promontory bone close to the OW (see text for details regarding BC velocity calculations). With the same stimulation in the same specimen, the difference velocity peak magnitude (c) and phase (d) of 30 positions on the RW membrane is shown [Fig. 2(b)]. The phase of the difference velocity in panels (b) and (d) are calculated relative to the stimulation velocity. The spatial resolution on the RW membrane is 0.2 mm and the frequency resolution of the measurements are 50 points/decade.

The average fluid volume displacement at the OW and RW of the eight temporal bone specimens, when a BC stimulation velocity of 1 mm/s was used, is shown in Fig. 9. It should be noted that the malleus, incus, and the stapes superstructure were removed when the BC measurements were made. This means that Fig. 9 does not present the normal results of stapes footplate and RW membrane motion with BC stimulation; the influence for the external ear and the middle ear ossicles was absent. However, the fluid flow at the RW and OW can still be compared. At low frequencies, the RW volume displacement is greater than the volume displacement at the OW; they are approximately equal at 3 kHz, while above this frequency, the volume displacement at the OW is greater than at the RW. The interindividual differences of fluid volume displacement were greater with BC stimulation than with AC stimulation; this is indicated by the standard deviation of the measurements (vertical bars in Figs. 5 and 9).

The phase of the fluid volume displacement at the OW in Fig. 9 is close to 90° for frequencies below 0.7 kHz; above

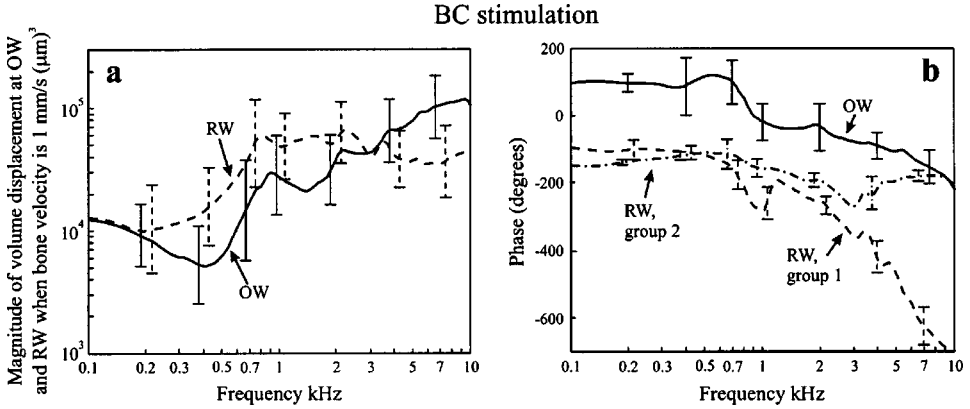


FIG. 9. Average peak magnitude in cubic micrometers (a) and phase (b) of the fluid volume displacement from eight temporal bone specimens. The specimens are stimulated with a BC vibration of 1 mm/s. The volume displacement at the OW is given by the solid line and the volume displacement at the RW by the dashed line. The phases of the volume displacements are related to the stimulation velocity of the temporal bone. For the OW, the phase is the average from eight specimens (solid line), whereas, for the RW, the phases are divided into two groups. Each group is the average of four specimens. Group 1 (dashed line) contains the specimens where the phase falls off at the higher frequencies, and group 2 (dashed-dotted line) contains the specimens with an almost flat phase. The vertical bars are the standard deviation of the volume displacement. The frequency resolution is 50 points/decade.

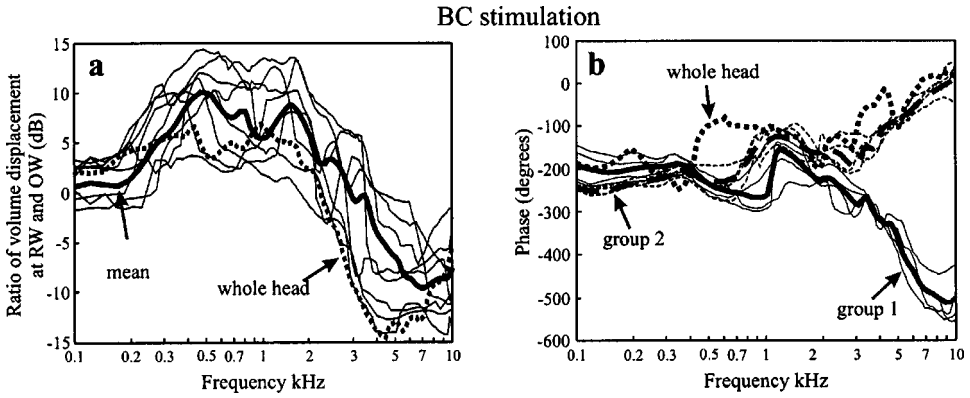


FIG. 10. Level (a) and phase (b) of the ratio of volume displacement at the RW and OW when stimulation is by BC. The data are calculated as the volume displacement at the RW divided by the volume displacement at the OW. The individual result of the eight specimens is plotted with thin lines and the average of the eight curves is plotted with a thick line (a). In the phase panel (b), the phase of the four specimens belonging to group 1 is plotted with thin solid lines and the average of those four is plotted with a thick solid line while the phase of the specimens belonging to group 2 is plotted with a thin dashed line and the average of those four is plotted with a thick dashed line. In addition, the level and phase of the volume displacement ratio for BC stimulation between the RW and the OW, from a human cadaver head, is included (dotted line). The frequency resolution is 50 points/decade.

0.7 kHz it falls off at approximately 80°/oct. The phase of the volume displacement at the RW is very different. When examining the individual phases for each specimen, it was found that they belonged to one of two categories: above 1 kHz, they had either a high slope of approximately −200°/oct (referred to as group 1) or were flattened (referred to as group 2). As can be seen from the standard deviations of the two groups (the vertical bars), the spread within each group is small. Of the eight specimens tested with BC stimulation, four belonged to each group. No anatomical difference, or any other difference, was found between the specimens belonging to the two groups.

With BC stimulation, there is a difference between the RW and OW volume displacement. The ratio of the volume displacement at the RW and OW is shown as level and phase in Fig. 10. Individual data from the specimens as well as the average is shown in Fig. 10. Although there is greater variability among the individual data, the trend is similar to the mean, both in level and phase. For the mean result in Fig. 10(a), 5 to 10 dB more fluid is displaced at the RW than at the OW between 0.2 and 2 kHz. Above 2 kHz, this difference diminishes at approximately −10 dB/oct and becomes flat at 7 kHz, with a difference of −10 dB. Also, the phase of the fluid flow between the two windows differs considerably. At the low frequencies, the phase difference is between 150° and 250°. For the specimens belonging to group 1, the phase difference falls off at a rate of −120°/oct above 1.2 kHz. For group 2 and at frequencies above 3 kHz, the phase difference rises by 140°/oct. This indicates clearly that the volume displacement at the two windows differs in both magnitude and phase.

Also included in Fig. 10 is the volume displacement measurement of the RW and OW with BC stimulation in a human cadaver head. The level difference between the RW and OW fluid displacements is similar for the cadaver head and the specimens. In a whole head and with BC stimulation, the phase between the fluid displacement of the two windows shows a tendency similar to the phases obtained in group 2 of the temporal bone specimens. The similarity of response between the whole head and temporal bone specimens for BC stimulation is taken as an indication that using temporal bone specimens did not bias the results.

C. Lesions of the cochlea

The alterations of the magnitude response of the RW fluid volume displacement with BC stimulation, after the stapes footplate was glued to the promontory bone, and also after a small hole (diameter 1 mm) was drilled into the stapes footplate, are given in Fig. 11. The result is presented as the RW volume displacement, when the stapes footplate is glued, relative to the RW volume displacement for a normal cochlea, at the same bone velocity. Similarly, the RW volume displacement with the stapes footplate fenestrated is also compared with a normal cochlea. When the stapes footplate is glued to the promontory bone, the relative velocity at the footplate is reduced more than 30 dB. This was verified with

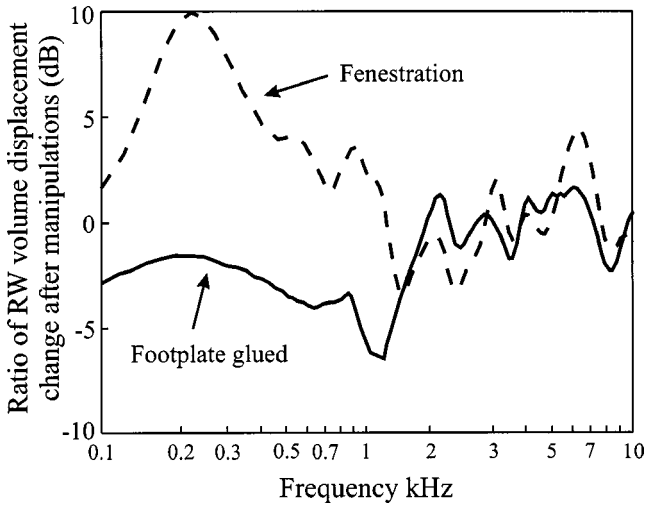


FIG. 11. The level change of the fluid flow at the RW after manipulations of the cochlea. The data are the average from the eight temporal bone specimens stimulated by BC. Solid line: volume displacement at the RW, with the stapes footplate glued to the otic capsule, compared with a normal intact cochlea. Dashed line: volume displacement at the RW with a hole of ∅ 1 mm drilled into the stapes footplate compared with a normal intact cochlea. The frequency resolution is 50 points/decade.

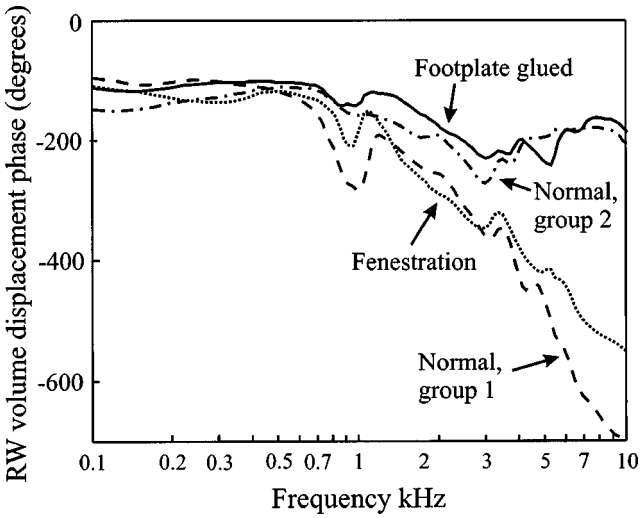


FIG. 12. The average phase of the RW volume displacement relative to the specimen stimulation velocity when the stimulation is by BC. Solid line: RW volume displacement phase after the stapes footplate was glued to the otic capsule, average of eight specimens. Dotted line: RW volume displacement phase after a hole of ∅ 1 mm was made in the stapes footplate, average of eight specimens. Dashed line: RW volume displacement phase of a normal intact cochlea (group 1), average of four specimens. Dashed-dotted line: RW volume displacement phase of a normal intact cochlea (group 2), average of four specimens. The frequency resolution is 50 points/decade.

LDV measurements of the footplate motion, after the footplate was fixed. However, with the stapes footplate glued, the fluid volume displacement at the RW is reduced only by 2 to 5 dB for frequencies below 1.5 kHz. Above 1.5 kHz, the average difference in the RW fluid volume displacement is less than 2 dB between a normal cochlea and one with the stapes footplate fixed. When a hole in the footplate is opened, the fluid volume displaced at the RW increases by up to 10 dB at 0.2 to 0.3 kHz. This difference diminishes with frequency and is close to 0 dB at 1 kHz. Above 1 kHz, the difference is less than 3 dB between the opened footplate and the normal condition.

Shown in Fig. 12 is the average phase of the fluid volume displacement at the RW for the three conditions: normal, footplate glued, and footplate fenestrated. Here, the phases of the volume displacements are related to the promontory bone velocity in the stimulation direction [same as in Fig. 9(b)]. The two normal conditions of the RW volume displacement (groups 1 and 2) are included for comparison with the phases obtained when the stapes footplate was glued or when it was fenestrated. When the stapes footplate is glued to the otic capsule, the phase of the RW volume displacement stays approximately between −100° and −200° for the frequency range 0.1 to 10 kHz. This phase is close to the RW volume displacement phase for the normal cochlea (group 2). This means that the RW volume displacement phases of the specimens belonging to group 2 remained similar when the footplate was glued, whereas the phases of the specimens belonging to group 1 altered considerably. All specimens showed similar RW volume displacement phases after the stapes footplate was glued to the otic capsule.

When a hole was made in the stapes footplate (fenestration), the phase of the RW volume displacement was not the same as when the footplate was glued. After the fenestration, the RW volume displacement phase remained between −100° and −150° at the low frequencies; above 1 kHz the phase started to fall off by about −130°/oct. This phase of the RW volume displacement is close to the normal RW volume displacement phase for group 1. In this case, the specimens from group 1 have, after a fenestration in the footplate, a phase of the RW volume displacement which is close to the original normal RW volume displacement phase. After the footplate was fenestrated, the specimens from group 2 exhibited quite different phases of the RW fluid volume displacement.

IV. DISCUSSION

A. Measurement method

The volume displacements at the RW and OW were obtained by integrating the displacements of small segments of the RW membrane and stapes footplate. This method offered no problem in the estimation of the volume displacement of the stapes footplate; the footplate was assumed to move as a stiff plate for frequencies up to 10 kHz, and the velocity measurements of the footplate verified this assumption. Since the RW is a membrane, the spatial resolution of the segments must be high with respect to the vibration modes of the membrane. This was not a problem at the low frequencies, where the membrane moved in a simple in-and-out motion. However, the number of vibration modes increased with frequency, and some bias could have affected the results at the highest frequencies used (above 8 kHz) due mainly to the uncertainty of the exact target position. The vibration modes of the RW membrane corresponded well with those reported in cats with time-averaged holography (Khanna and Tonndorf, 1971; Nomura, 1984).

The vibration mode of the specimens with BC stimulation was controlled; it constituted more or less translational vibration in one direction. This is different from normal BC stimulation in a human head. The temporal bone of an intact skull with BC stimulation moves in all three dimensions, without any dominating vibration direction (Stenfelt *et al.*, 2000). Hence, the volume displacement in a normal skull could differ considerably from the controlled measurements conducted here. However, the fluid volume displacement results for the human head with BC stimulation were similar to the average data from the specimens with BC stimulation (Fig. 10). It is therefore unlikely that the mode of stimulation affected the results of the volume displacement when stimulation was by BC.

Another concern is that air bubbles may have entered the cochlea. An air bubble in the cochlea causes a compliance that affects the fluid flow at the RW and OW with AC as well as BC stimulation. Authors pointing out problems with air bubbles in the cochlea, e.g., Puria *et al.* (1997), have opened the cochlear space in their experiments. That was not done in this study. The inner ear and its fluid canals (e.g., vestibular and cochlear aqueducts) were left intact. The temporal bone bones were stored in saline from the time of autopsy to the experiment time. During the experiment, sealing of the specimen in latex and spraying the specimen with saline were used to prevent drying that could, in time, lead to air

bubbles in the cochlea. Further, with AC stimulation, an air bubble would shunt the fluid flow at the RW and cause a difference in volume displacement at the two windows; this was not seen in the result. Moreover, similar preparation technique was conducted in previous reports, e.g., Hato *et al.* (2003) and Stenfelt *et al.* (2002), where no air bubbles were found in the cochleae. Since the results from the whole head (Fig. 10), where no preparation was done of the inner ear part, showed similar result as the temporal bone specimens, air bubbles in the cochlea are not believed to influence the results found here.

B. AC measurements

A general assumption is that the cochlear contents are incompressible and, therefore, the volume flow at the two windows (RW and OW) should be equal and of opposite phases. Figures 5 and 6 verify this assumption when stimulation is by AC. The maximum deviation in the average fluid flow between the two windows below 1 kHz is less than 3 dB, and, for frequencies above 1 kHz, the difference between the average fluid volume displacements at the two windows is close to 0 dB. The phase between the volume displacements at the two windows is close to 180° for nearly the whole frequency range measured; it is only at the very high frequencies, above 7 kHz, that the phase difference diminishes. The volume displacement at the RW is slightly greater than the data reported by Kringlebotn and Gundersen (1985). However, the displacement of the stapes footplate is similar to previous reports (Aibara *et al.*, 2001).

In a study of the volume displacement at the two windows in pig temporal bone specimens, Kringlebotn (1995) reported the volume flows to be within 1 dB for frequencies below 2.4 kHz. In the same study, a measurement was conducted on a single human temporal bone specimen. The volume flow between the windows was found to be within 5 dB for frequencies below 2 kHz; this volume difference was attributed to uncertainties in the measurements. For frequencies below 2 kHz, those results are similar to those presented here with AC stimulation. In a measurement of the cochlear potentials in cats, Wever *et al.* (1948) showed that similar results were obtained regardless of whether the tone was applied at the OW or the RW. This indicates indirectly that the OW and RW have equal fluid flow for a specific stimulation. In another study of cochlear potentials in cats, Wever and Lawrence (1950) applied a tone of equal magnitude but different phase to the windows. At a certain phase relation, the cochlear potential was twice that of a single stimulation, and a minimum was found at a phase relation 180° from this setting. This is yet another indirect indication that both windows have equal fluid flow. However, the absolute phase was not reported in the test of Wever and Lawrence.

Tonndorf and Tabor (1962) found 20 to 30 dB reduction in cochlear potentials in cats after the RW was blocked. Again, inadequate blocking of the RW was given as the source for the fluid flow. However, with a moderate to severe hearing loss, patients with RW otosclerosis are still able to hear (Huygen *et al.*, 1974). Also, persons without a RW still have some hearing ability. This hearing was improved by a RW fenestration operation (Harrison *et al.*, 1964; Richards, 1981). In these cases, even if no fluid flow were possible at the RW, a traveling wave on the basilar membrane was present, although with a reduction of at least 20 to 30 dB compared with the normal intact ear. This indicates that there are some compliant structures in the cochlea that allow a fluid flow in the event when no outlet is possible at the RW. However, this compliant structure has at least 20 to 30 dB greater impedance than the RW and does not affect the fluid flow in a normal intact cochlea.

Peake *et al.* (1992) derived a model based on the assumption that the stimulation of the cochlea was due to the pressure difference between the two cochlear windows. For frequencies above 0.8 kHz, this model predicted results similar to previously obtained average hearing results. This supported their conclusion that the dominant stimulation of the cochlea was the sound pressure difference between the two windows. In an investigation to determine whether the stimulation mode of the cochlea was a pressure difference between the two windows, or if the cochlea is sensitive to a common mode stimulation at the two windows (same sound pressure at both windows), Voss *et al.* (1996) concluded that, in a normal ear, the main stimulation mode was by a differential sound pressure at the windows. These investigations showed further that the fluid flow at the two windows in a healthy cochlea should be equal and of opposite phase.

C. BC measurements

The results in Figs. 9 and 10 showed that, with BC stimulation, the fluid volume displacement at the two windows is not equal and of opposite phase. This is even more obvious when there is a lesion of the stapes footplate (Figs. 11 and 12). That the fluid flow was unequal at the two windows in a cochlea with lesions was expected. Carhart (1962) reports mainly a decrease in the BC sensitivity around 2 kHz with otosclerosis of the stapes footplate; at the lower and higher frequencies, the BC sensitivity was only slightly affected. After a surgical stapes mobilization, fenestration of the cochlea, or total footplate stapedectomy, the BC thresholds generally improved and became close to normal (Walsh, 1962; Awengen, 1993). Everberg (1968) found that BC thresholds were close to normal in patients with no OW due to malformations. Martin *et al.* (2002) and Linder *et al.* (2003) reported BC hearing close to normal in patients with no RW. Tonndorf and Tabor (1962) occluded the RW in cats and found, by measuring the cochlear microphonics, no significant change in the BC sensitivity.

With BC stimulation, the phases of the volume displacement at the OW were similar for the eight specimens tested, whereas the phases of the volume displacement at the RW belonged to one of two groups: it either fell off at high frequencies or remained almost constant over the whole measured frequency range (Fig. 9). No explanation for this could be found in the preparation of the specimens, the attachment of the shaker, or the measurement method. These were all the same for all eight specimens. No detailed analysis of the anatomy of each specimen was conducted, since the difference in phase among the specimens was not discovered until the post analysis of the data; the specimens had been destroyed by then. Our hypothesis is that, with BC stimulation,

additional fluid flow is provided by one or a combination of the following: (1) compression and expansion of the cochlear bone, which alters the volume of the cochlear space; (2) small compressibility of the cochlear fluid; and (3) the small canals and compliant structures of the cochlea [termed third window by Ranke *et al.* (1952)]. This additional fluid flow is superimposed in phase or out of phase with the fluid flow in the cochlea, which results in the two phases seen. However, the reason for the in phase or out of phase addition is unknown to us.

The artificial lesions of the stapes footplate provide a hint to the mystery of the different phases. When the stapes footplate was glued to the annulus, the impedance at the OW increased and the phase at the RW remained relatively constant (Fig. 12). When a hole was drilled into the stapes footplate, the impedance at the OW decreased and the phase at the RW fell off at higher frequencies. However, the impedance at the OW seems to affect only the displacement level at the RW for low frequencies. With a high impedance at the OW (footplate glued), the level of RW displacement was lowered by 2 to 5 dB below 2 kHz; above 2 kHz the difference was close to 0 dB (Fig. 11). This can be explained by the fact that no fluid is displaced at the OW, and the other phenomena causing a fluid flow in the cochlea are not equally efficient at the low frequencies; they are, however, equally effective above 2 kHz. When a hole was drilled into the stapes footplate, the level of fluid displaced at the RW increased by about 10 dB at the very low frequencies; this difference diminished with frequency and, for frequencies above 1 kHz, there was no difference at all. This can be explained by the stiffness of the annular ligament at the OW, which causes high impedance at the low frequencies [estimated to be about 20 times higher than the stiffness of the RW (Kirikae, 1959)]. With a hole in the stapes footplate, this stiffness was removed, which led to a greater fluid flow at the low frequencies.

D. Compressibility of the cochlea and the cochlear fluid

One of the very first explanations for BC sound was that the bone surrounding the cochlea, due to the vibrations of the bone, expands and compresses (Herzog, 1926; Krainz, 1926). This would lead to a compression and expansion of the cochlear space, which would cause a fluid flow in the cochlea. Tonndorf (1966) termed this mode distortional BC. The volume displaced at the windows, with a velocity of the bone of 1 mm/s, was somewhere between 10^4 and 10^5 cubic micrometers. Wever and Lawrence (1954) estimated the fluid volume of the cochlea to be 98.1 cubic millimeters. Hence, the volume displaced at the OW and RW, in a specimen with a bone velocity of 1 mm/s, is one-millionth (10^{-6}) of the total volume in the cochlea. For comparison, with an AC stimulation of 80 dB SPL at the TM, the displaced volume at the windows is about half of that volume. This means that only a very small compressibility of the cochlear fluid and cochlear bone is required to obtain this result.

The specimens used did not have any wave motion; they were vibrating as a single body throughout the whole frequency range used. This means that there was no expansion and compression of the bone due to distortional BC. However, the cochlear fluid is not completely incompressible, and the cochlear bone is not infinitely stiff. Hence, there can be some fluid flow in the cochlea, even if fluid inertia is the only mode of stimulation. Shera and Zweig (1992) concluded that a small compressibility of the scala media could explain results of low frequency hearing without a TM and middle ear ossicles, reported by Békésy (1936). Wever and Lawrence (1954) stated that the compressibility of water was, when recalculated to SI units, 4.9×10^{-12} at a pressure of 1 Pa. Hence, if compression of the fluid in the cochlea were the only explanation for the volume difference at the windows, the pressure in the cochlea at a bone velocity of 1 mm/s needs to be more than 10^5 Pa. Clearly, this is not the case. They stated further that the yielding of the cochlear bone was even less than that of the fluid. However, no proof for that statement was given.

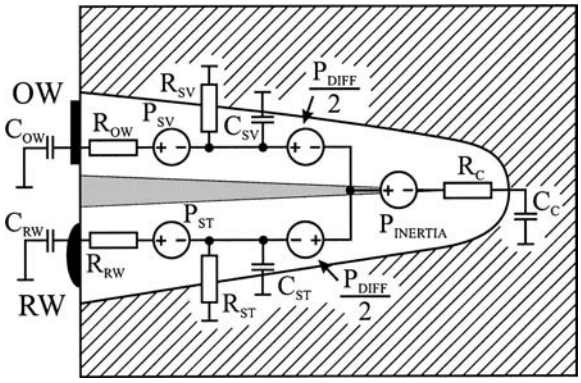


FIG. 13. The figure describes a simplified lumped parameter model of fluid inertia in the cochlea when stimulation is by BC. The physical interpretation of the components is given in the text.

To facilitate grasping the functions in the cochlea under the influence of inertial forces in the fluid, due to BC stimulation, a simplified model of the cochlea including lumped parameters is presented in Fig. 13. The cochlea with scala vestibule (SV) and scala tympani (ST) is depicted together with the RW and OW. There are four sources contributing to pressure in the cochlear fluids (as measured in front of each window). $P_{INERTIA}$ is the general common-mode pressure in the fluid at the two windows, whereas P_{DIFF} is the pressure difference caused by a pressure gradient between the two windows; P_{SV} and P_{ST} are the inertial pressures caused by the fluid close to the windows in the scala vestibuli and scala tympani, respectively. These two sources, which are in phase, are determined by local effects near the windows. The $P_{INERTIA}$ is mainly influenced by the acceleration of the bone surrounding the cochlea and the mass of the fluid, whereas the P_{DIFF} is mainly influenced by the geometry of the cochlea and the vibration directions which produces a pressure gradient between the two windows. The impedances of the RW and OW are modeled by the compliances C_{RW} and C_{OW}, respectively, while R_{RW} and R_{OW} represent the loss due to the viscosity in the fluid close to the windows. The ducts enabling fluid flow on the scala tympani side close to the RW are modeled with R_{ST} and, in a similar way, R_{SV} represents the ducts on the scala vestibuli side, close to the OW. The

C_{ST} represents the compliance of the fluid and cochlear bone on the scala tympani side close to the RW, whereas C_{SV} represents the compliance of the fluid and cochlear bone on the scala vestibuli side close to the OW. The compliance of the major part of the fluid volume in the cochlea, together with the compliance of the cochlear bone and other compliant structures, such as nerves or blood vessels entering the cochlea, is designated C_C. Finally, R_C models the damping of the fluid, due mainly to fluid viscosity.

Although no values of the parameters are given in Fig. 13, the model shows that, provided the impedance magnitudes are not too high compared with the inertial pressures, inertial effects of the fluid can produce a fluid flow in the cochlea with BC stimulation, even if the OW, the RW, or both, are obstructed. It should be noted that the model in Fig. 13 is derived for BC inertial fluid stimulation. Any other type of stimulation, AC as well as BC, would produce a different model.

E. Aqueducts and compliant structures of the cochlea

The aqueducts into the cochlea are believed to influence the hearing of BC sound, at least when lesions affect the ear. Communication between the cochlea and the cranial cavity was termed the third window by Ranke *et al.* (1952). This includes fluid flow through the perilymphatic or endolymphatic ducts, as well as flow produced by compression of cochlear structures such as nerves, blood vessels, or scala media. Micro-channels have been identified that connect the cochlea with surrounding structures (Kucuk *et al.*, 1991). These micro-channels may also contribute to the fluid flow. The ducts of the cochlea, which are long and narrow, offer high impedance to fluid flow. Ranke (1958) suggested that all possible communications between the inner ear and cranial space should be regarded as potential pathways for BC stimulation. It was shown by Freeman *et al.* (2000) and Sohmer *et al.* (2000) that the coupling between the cranial space and the cochlea transmits hydrodynamic pressures. They found that an intracranial pressure gave a hearing sensation greater than could be explained by the vibrations of the skull bone alone. Schneider (1959) opposed the internal auditory canal as a transmission path for sound pressure; he did not comment on any other possible transmission.

It has been shown that a static pressure in the cerebrospinal fluid is transmitted to the cochlear fluids (Yoshida and Uemura, 1991), mainly through the cochlear aqueduct. Gopen *et al.* (1997) showed, with a model, that the impedance of the cochlear aqueduct is too high to influence the hearing in a healthy ear, at least with AC stimulation. Groen and Hoogland (1958) found, in a patient with an obstructed RW, a deterioration of the BC sensitivity at the high frequencies of 6 dB/oct; they attributed this to the impedance of the aqueduct. However, Tonndorf and Tabor (1962) reported no significant change of BC sensitivity in cats, after occluding the RW provided an open cochlear aqueduct. Békésy (1936) suggested that, since ears without a TM and middle ear ossicles had a low frequency response to AC stimulation, a volume displacement may arise through the endolymphatic or paralymphatic ducts, or through the cochlear blood vessels.

A patent cochlear and vestibular aqueduct could, at least partly, explain the fluid flow at the OW and RW found for AC and BC stimulation. On the one hand, in a normal ear stimulated by AC sound, the impedance of the aqueducts is much larger than the impedance of the RW. Therefore, almost no fluid is displaced in the aqueducts, and the fluid displaced at the RW equals that at the OW, but with opposite phase. For BC stimulation, on the other hand, the acceleration of the fluid in the cochlea causes a pressure distribution in the fluid. This pressure distribution gives a fluid flow, not only at the RW and OW, but also through the aqueducts. Due to this pressure distribution, the pressure between the aqueducts and one of the windows (P_{SV} or P_{ST} in Fig. 13) can be greater than the pressure between the OW and the RW. Consequently, some of the fluid in the cochlea can be displaced through the aqueducts, and the fluid flows at the OW and RW differ in magnitude. Also, the phase between the fluid flows at the two windows differs from the expected 180°. An explanation could be that, since temporal bone specimens are used, the aqueducts were severed in the extraction process which caused the effect observed. However, care was taken when removing the temporal bones from the heads not to damage any of the cochlear structure. Further, similar results were obtained when the whole skull was used. This indicates that the result with different fluid flow at the RW and OW is not due to damage of the aqueducts communicating fluid to the cochlea.

The hypothesis that the difference in fluid flow between the two windows, for BC stimulation, originates in compression and expansion of the cochlear bone, compression and expansion of the fluid itself, from fluid flow through the small channels of the cochlea, and from the yielding of compliant structures in the cochlea, could not be verified. However, a plausible explanation is a combination of the abovementioned phenomena. At low frequencies, the impedance of fluid flow through the aqueducts is low, whereas the stiffness-controlled impedances, such as compression of the bone surrounding the cochlea and nerves entering the cochlea, are of high magnitudes. At the high frequencies, this is reversed: the aqueducts have high impedance for fluid flow, whereas stiffness-controlled phenomena offer lower impedance. If the results are a combination of these different effects, this could explain the differences in phases and magnitudes of the volume flow at the two windows when stimulation is by BC.

V. CONCLUSIONS

The fluid volume displacement at the RW and the OW was estimated by LDV measurements in temporal bone specimens. The velocity at five positions on the stapes footplate was used to estimate the volume displacement at the OW, while the velocity at up to 40 positions on the RW membrane, with a spacing of 0.2 mm, was measured to obtain the volume displacement at the RW.

When the temporal bone specimens were stimulated by AC sound, the average fluid volume displacements at the RW and OW were within 3 dB of each other for frequencies

below 1 kHz; above 1 kHz, the fluid volumes displaced at the two windows were approximately equal. The phase difference of the fluid flow at the RW and OW was close to 180° for the frequency range measured, except above 7 kHz where the difference diminished.

For BC stimulation, the volume displacement at the two windows differed by 5 to 15 dB for frequencies below 2 kHz, with more fluid displaced at the RW. Above this frequency, the difference changed with frequency and, at 7 kHz, 5 to 15 dB more fluid was displaced at the OW. The phase difference of the fluid volume displacement at the two windows, for BC stimulation, falls into one of two groups. At the low frequencies, the phase between the fluid flows at the two windows is 150° to 250°; above 2 kHz, the phase either rises by about 80°/oct or falls by −100°/oct. Measurements in an intact human head gave result similar to the temporal bone specimens with BC stimulation.

When the stapes footplate was glued to the otic capsule, and the specimen was stimulated by BC, the RW fluid volume displacement was almost the same as in a normal intact specimen; only a small deterioration of 2 to 5 dB below 1 kHz was seen. When a small hole (diameter 1 mm) was made in the stapes footplate, the RW fluid volume displacement with BC stimulation increased by up to 10 dB for frequencies below 1 kHz; above this frequency no significant change was seen.

The results of unequal fluid displacement at the two windows, for BC stimulation, were hypothesized to be a result of several phenomena: compression and expansion of the cochlear bone, compression and expansion of the fluid itself, fluid flow through the small channels of the cochlea, and yielding of compliant structures in the cochlea. However, it has not yet been possible to verify these phenomena.

ACKNOWLEDGMENTS

This work was supported in part by a V. A. Merit Review Grant (GDE0010ARG), the Swedish Institute, and Swedish Research Council for Engineering Sciences (TFR 299-2000-576).

Aibara, R., Welch, J., Puria, S., and Goode, R. (**2001**). ''Human middle ear sound transfer function and cochlear input impedance,'' Hear. Res. **152**, 100–109.

Awengen, D. (**1993**). ''Change of bone conduction thresholds by total footplate stapedectomy in relation to age,'' Am. J. Otolaryngol. **14**, 105–110.

von Békésy, G. (**1936**). ''Zur Physik des Mittelohres und über das Hören bei fehlerhaftem Trommelfell (Physics of the middle ear and on hearing with the eardrum absent),'' Akust. Z. **1**, 13–23.

Carhart, R. (**1962**). ''Effect of stapes fixation on bone conduction response,'' in *Otosclerosis*, edited by H. Schuknecht (Little, Brown, Boston), pp. 175–197.

Carlborg, B., Densert, B., and Densert, O. (**1982**). ''Functional patency of the cochlear aqueduct,'' Ann. Otol. **91**, 209–215.

Everberg, G. (**1968**). ''Congenital absence of the oval window,'' Acta Otolaryngol. (Stockh) **66**, 320–332.

Freeman, S., Sichel, J.-Y., and Sohmer, H. (**2000**). ''Bone conduction experiments in animals—evidence for a non-osseous mechanism,'' Hear. Res. **146**, 72–80.

Gopen, Q., Rosowski, J., and Merchant, S. (**1997**). ''Anatomy of the normal human cochlear aqueduct with functional implications,'' Hear. Res. **107**, 9–22.

Groen, J. J., and Hoogland, G. A. (**1958**). ''Bone conduction and otosclerosis of the round window,'' Acta Otolaryngol. (Stockh) **49**, 206–212.

Håkansson, B., Brandt, A., Carlsson, P., and Tjellström, A. (**1994**). ''Resonance frequency of the human skull *in vivo*,'' J. Acoust. Soc. Am. **95**, 1474–1481.

Harrison, W., Shambaugh, G., and Derlacki, E. (**1964**). ''Congenital absence of the round window: Case report with surgical reconstruction by cochlear fenestration,'' Laryngoscope **74**, 967–978.

Hato, N., Stenfelt, S., and Goode, R. (**2003**). ''Three-dimensional stapes footplate motion in human temporal bones,'' Audiol. Neuro-Otol. **8**, 140–152.

Herzog, H. (**1926**). ''Das Knochenleitungsproblem. Theoretische Erwagungen,'' Z. Hals Nas. Ohrenheilk. **15**, 300–306.

Huygen, P. L., Marres, E. H., and Brand, D. J. (**1974**). ''Focus localization in two clinical types of otosclerosis,'' Acta Otolaryngol. (Stockh) **78**, 365–370.

Khanna, S., and Tonndorf, J. (**1971**). ''The vibratory pattern of the round window in cats,'' J. Acoust. Soc. Am. **50**, 1475–1483.

Kirikae, I. (**1959**). ''An experimental study on the fundamental mechanism of bone conduction,'' Acta Otolaryngol. Suppl. (Stockh) **145**, 1–110.

Krainz, W. (**1926**). ''Das Knochenleitungsproblem. Experimentelle Ergebnisse,'' Z. Hals Nas. Ohrenheilk. **15**, 306–313.

Kringlebotn, M. (**1995**). ''The equality of volume displacement in the inner ear windows,'' J. Acoust. Soc. Am. **98**, 192–196.

Kringlebotn, M., and Gundersen, T. (**1985**). ''Frequency characteristics of the middle ear,'' J. Acoust. Soc. Am. **77**, 159–164.

Kucuk, B., Abe, K., Ushiki, T., Inuyama, Y., Fukuda, S., and Ishikawa, K. (**1991**). ''Microstructures of the bony modiolus in the human cochlea: A scanning electron microscopic study,'' J. Electron Microsc. **40**, 193–197.

Linder, T. E., Ma, F., and Huber, A. (**2003**). ''Round window atresia and its effect on sound transmission,'' Otol. Neurotol. **24**, 259–263.

Martin, C., Tringali, S., Bertholon, P., Pouget, J.-F., and Prades, J.-H. (**2002**). ''Isolated congenital round window absence,'' Ann. Otol. Laryngol. **111**, 799–801.

Nomura, Y. (**1984**). ''Otological significance of the round window,'' Adv. Oto-Rhino-Laryngol. **33**, 1–162.

Palva, T., and Dampert, K. (**1969**). ''Human cochlear aqueduct,'' Acta Otolaryngol. Suppl. (Stockh) **246**, 1–58.

Peake, W., Rosowski, J., and Lynch, III, T. (**1992**). ''Middle-ear transmission: Acoustic versus ossicular coupling in cat and human,'' Hear. Res. **57**, 245–268.

Puria, S., Peake, W., and Rosowski, J. (**1997**). ''Sound-pressure measurements in the cochlear vestibule of human cadaver ears,'' J. Acoust. Soc. Am. **101**, 2754–2770.

Ranke, O. (**1958**). ''Discussion remark to Meyer z. Gottesberge, A.: Die Schalleitung im Mittelohr in klinischer Sicht,'' Z. Laryngol. **37**, 355–367.

Ranke, O., Keidel, W., and Weschke, H. (**1952**). ''Des Hören beim Verschluss des runden Fensters,'' Z. Laryngol. **31**, 467–475.

Rask-Andersen, H., Stahle, J., and Wilbrand, H. (**1977**). ''Human cochlear aqueduct and its accessory canals,'' Ann. Otol. Rhinol. Laryngol. Suppl. **86**, 1–16.

Richards, S. H. (**1981**). ''Congenital absence of the round window treated by cochlear fenestration,'' Clin. Otolaryngol. **6**, 265–269.

Ritter, F. N., and Lawrenece, M. (**1965**). ''A histological and experimental study of cochlear aqueduct patency in the adult human,'' Laryngoscope **75**, 1224–1233.

Schneider, W. (**1959**). ''Gegenbeweis gegen Knochenleitung mittels Druckwellen über den Kanal des Nervus acusticus,'' Laryngol. **11**, 723–734.

Shera, C., and Zweig, G. (**1992**). ''An empirical bound on the compressibility of the cochlea,'' J. Acoust. Soc. Am. **92**, 1382–1388.

Sohmer, H., Freeman, S., Geal-Dor, M., Adelman, C., and Savion, I. (**2000**). ''Bone conduction experiments in humans—a fluid pathway from bone to ear,'' Hear. Res. **146**, 81–88.

Stenfelt, S., Håkansson, B., and Tjellström, A. (**2000**). ''Vibration characteristics of bone conducted sound *in vitro*,'' J. Acoust. Soc. Am. **107**, 422–431.

Stenfelt, S., Hato, N., and Goode, R. L. (**2002**). ''Factors contributing to bone-conduction: The middle ear,'' J. Acoust. Soc. Am. **111**, 947–959.

Tonndorf, J. (**1962**). ''Experimental occlusion of both cochlear windows,'' in *Otosclerosis*, edited by H. Schuknecht (Little, Brown, Boston), pp. 119–126.

Tonndorf, J. (**1966**). ''Bone conduction. Studies in experimental animals,'' Acta Otolaryngol. Suppl. (Stockh) **213**, 1–132.

Tonndorf, J., and Tabor, J. (**1962**). ''Closure of the cochlear windows: Its effect upon air and bone conduction,'' Ann. Otol. Rhinol. Laryngol. **71**, 5–29.

Voss, S., Rosowski, J., and Peake, W. (**1996**). ''Is the pressure difference between the oval and round windows the effective acoustic stimulus for the cochlea,'' J. Acoust. Soc. Am. **100**, 1602–1616.
Walsh, T. (**1962**). ''Fenestration: Results, indications, limitations,'' in *Otosclerosis*, edited by H. Schuknecht (Little, Brown, Boston), pp. 245–250.
Wever, G., and Lawrence, M. (**1950**). ''The acoustic pathways to the cochlea,'' J. Acoust. Soc. Am. **22**, 460–467.
Wever, E. G., and Lawrence, M. (**1954**). *Physiological Acoustics* (Princeton U.P., Princeton), p. 454.
Wever, G., Lawrence, M., and Smith, K. (**1948**). ''The middle ear in sound conduction,'' Arch. Otolaryngol. **48**, 19–35.
Yoshida, M., and Uemura, T. (**1991**). ''Transmission of cerebrospinal fluid pressure changes to the inner ear and its effect on cochlear microphonics,'' Eur. Arch. Otorhinolaryngol. **248**, 139–143.

A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception

Sophie K. Scott[a)]
Departments of Psychology and Phonetics & Linguistics, University College London, London WC1E 6BT, United Kingdom

Stuart Rosen
Department of Phonetics & Linguistics, University College London, London WC1E 6BT, United Kingdom

Lindsay Wickham
Department of Psychology, University College London, London WC1E 6BT, United Kingdom

Richard J. S. Wise
MRC Clinical Sciences Centre, London W12 0NN, United Kingdom

(Received 27 September 2002; revised 14 November 2003; accepted 17 November 2003)

Positron emission tomography (PET) was used to investigate the neural basis of the comprehension of speech in unmodulated noise (''energetic'' masking, dominated by effects at the auditory periphery), and when presented with another speaker (''informational'' masking, dominated by more central effects). Each type of signal was presented at four different signal-to-noise ratios (SNRs) (+3, 0, −3, −6 dB for the speech-in-speech, +6, +3, 0, −3 dB for the speech-in-noise), with listeners instructed to listen for meaning to the target speaker. Consistent with behavioral studies, there was SNR-dependent activation associated with the comprehension of speech in noise, with no SNR-dependent activity for the comprehension of speech-in-speech (at low or negative SNRs). There was, in addition, activation in bilateral superior temporal gyri which was associated with the informational masking condition. The extent to which this activation of classical ''speech'' areas of the temporal lobes might delineate the neural basis of the informational masking is considered, as is the relationship of these findings to the interfering effects of unattended speech and sound on more explicit working memory tasks. This study is a novel demonstration of candidate neural systems involved in the perception of speech in noisy environments, and of the processing of multiple speakers in the dorso-lateral temporal lobes. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1639336]

PACS numbers: 43.64.Sj [PFA] Pages: 813–821

I. INTRODUCTION

Conversing at a cocktail party has been used as a classic demonstration of how we can listen to one person while ignoring the surrounding babble, yet also detect relevant spoken information such as one's name (Cherry, 1953; Conway *et al.*, 2001; Moray, 1959). This apparent processing of ignored auditory information is consistent with evidence that irrelevant speech signals can disrupt performance in auditory-verbal tasks, such as those involving verbal working memory (Tremblay *et al.* 2000). In trying to understand at least some aspects of the ''cocktail party effect,'' a distinction has been made between ''energetic'' and ''informational'' masking [see Brungart (2001) for a review].

''Energetic'' masking is demonstrated most clearly when a speech signal is presented together with a steady-state wideband noise. Here, the interfering effects of the masker arise primarily in the cochlea, reflecting the frequency analysis exacted on the basilar membrane. Just as for the energetic masking of tones by noise, the degree of energetic masking of speech by noise is primarily determined by the frequency spectrum of the noise, and its relative intensity (assumptions incorporated in articulation index theory) which computes SNR values within frequency channels (French and Steinberg, 1947). In particular, the lower the signal-to-noise ratio (SNR), the greater the masking of the signal.

''Informational'' masking, on the other hand, is supposed to reflect the interference of a masker on a target signal due to the masker having similar (and perhaps competing) informational content (Dirks and Bower, 1968; Festen and Plomp, 1990). The prime example of this is when speech is used to mask speech. Of course, energetic masking must also occur in such situations, though the amplitude modulations of natural speech alone impair its effectiveness as an energetic masker. The degree to which informational masking differs from energetic masking is demonstrated by the finding that the intelligibility of attended speech masked by other speech is relatively constant over a range of SNRs between −12 and 0 dB, and increases with increasingly positive SNRs (Brungardt, 2001). This level independency may arise from other sources, including the possibility that the amplitude modulations of a masking voice may afford ''glimpses'' of the target signal (Festen and Plomp, 1990). Such glimpses would allow the perception of the target signal during modulations of the masker, and reduce the effects of overall masking speech level.

a)Electronic mail: sophie.scott@ucl.ac.uk

The aim of this study was to investigate the neural basis of these masking effects using a functional neuroimaging technique [positron emission tomography (PET)], which tracks neural activity in terms of regional cerebral blood flow changes. In particular, we wanted to determine the neural correlates of the perception of speech masked by different signals, and also to determine the extent of SNR-dependent neural responses within this. The maskers chosen were speech and steady-state noise (with the same long-term spectral shape as speech). These two maskers were chosen as clear exemplars of informational and energetic masking, respectively. Within each condition, four SNR levels were selected, such that performance differences in the profile of behavioral responses (comprehension of sentences spoken by the target speaker) were clear between the speech and noise maskers without overall intelligibility falling below 60%. [See Davis and Johnsrude (2003) for an example of using noise masking as a method to manipulate intelligibility over a wider range of intelligibility levels.] The aim was also to use enough SNR levels, over a sufficient range, to maximize the possibility that neural correlates of level-dependent effects could be identified. To satisfy these constraints different SNR levels were chosen for each condition: -6, -3, 0, and $+3$ dB for the speech masker condition, and -3, 0, $+3$ and $+6$ dB for the noise masker condition. Pilot testing suggested that a SNR of -6 for the noise masker reduced intelligibility too much, while using a $+6$ dB SNR for the speech masker led to ceiling effects on the sentence comprehension task [consistent with positive SNRs being associated with an increase in intelligibility (Brungart 2001)]. The masking speech need not be intelligible to lead to interference effects in masking: studies have shown that reversed speech can also be an effective masker (Brungart and Simpson, 2002).

PET has several advantages for such a study, since it is *relatively* quiet compared to fMRI, and it does not suffer from signal loss due to susceptibility artefacts and geometric distortion (Devlin *et al.*, 2000). Susceptibility artefacts can be particularly problematic in the anterior temporal lobes, which we have previously demonstrated to be important in speech processing (Scott *et al.*, 2000). However, a disadvantage of PET is that the number of possible scans is limited by the total amount of radioactivity that can be administered. This study is thus preliminary in the sense that extensive testing of different masking conditions and different levels was not possible. Most importantly, this design did not permit the presentation of a baseline condition of listening to speech with no masking signal. In partial control for this, the results were contrasted with a previous study (Mummery *et al.*, 1999), in which passive speech perception [relative to signal correlated noise (Schroeder, 1969)] was studied.

There are several candidate neural systems that might be recruited during the attentional control of speech perception in energetic and informational masking. There could be a modulation of activation in primary auditory cortex (PAC) by different listening contexts, as reported by Ulanovsky *et al.* (2003). There could also be an alteration of the profile of activity in auditory association cortex, potentially linked to functional subsystems within this. For example, Griffiths and Warren (2002) have proposed that the planum temporale (posterior to PAC) operates as an informational hub for incoming auditory information, and that this might be associated with a distinct role of informational masking. Zatorre and colleagues (2002) have also emphasized the role of posterior auditory fields in the spatial representation of auditory objects. In contrast (though not necessarily contradiction) a meta-analysis of functional imaging studies has identified a role for the superior temporal gyrus (STG), lateral to primary auditory cortex (PAC), which is important in the processing of different aspects of auditory structure (e.g., AM, FM, harmonic structure), and which forms part of the acoustic processing of the speech signal (Scott and Johnsrude, 2003). Important in normal speech perception, this might be a candidate for the parallel processing of acoustic cues that are important for tracking a target voice and a masking signal in an informational masking context. There is thus the possibility that informational masking might be associated with a modulation of activity in the lateral STG and regions anterior to this, if informational masking results from competition between the processing of the target and the unattended voice within this system. Considering regions outwith the auditory system, it is possible that more generic, amodal attentional mechanisms could be recruited when speech is presented in a masking context. This would be associated with activation in prefrontal and parietal regions, commonly seen in cognitive tasks requiring the control of attention across modalities.

II. METHOD: SIGNAL PROCESSING AND STIMULI

All stimulus materials were drawn from digital representations (sampled originally at 22.05 kHz) of simple sentences recorded in an anechoic chamber by one male and one female speaker of standard Southern British English. The target sentences were always BKB sentences spoken by the female speaker whereas maskers were based on the IHR ASL sentences spoken by the male speaker (more details are given in the next section). All sentences were low-pass filtered at 3.8 kHz (sixth-order elliptical filter, both forward and backwards, so as to ensure zero-phase filtering equivalent to a 12th-order filter), and then downsampled to 11.025 kHz to save space. The masker and target signals were played together diotically (target and masker summed together and presented to both ears).

In the speech-in-noise condition, the target speech was played together with unmodulated noise, with the same long-term average spectrum as the masking male speaker. These calculations began with a spectral analysis of all 270 masker sentences, sampled at 22.05 kHz. Analyses used a FFT of length 512 sample points (23.22 ms), with windows overlapping by 256 points, giving a value for the spectrum at multiples of 43.1 Hz. The spectrum was then smoothed (in the frequency domain) with a 27-point Hamming window that was two octaves wide, over the frequency range 50 Hz to 7 kHz. The smoothed spectrum was then used to construct an amplitude spectrum for an inverse FFT (component phases randomized with a uniform distribution over the range $0-2\pi$) in order to create the speech-shaped noise.

Different SNRs were determined by a simple rms calculation across the entire waveform (e.g., target sentence), and

all combined waves were normalized to the same rms value.

The target and speech-masker stimuli were the BKB and IHR sentences respectively (Foster *et al.*, 1993; MacLeod and Summerfield, 1987). These are sets of syntactically simple sentences used to test intelligibility; they are scored according to the number of key words (two to three words per sentence) that are repeated correctly (e.g., ''she's brushing her hair,'' ''the clown had a funny face,'' ''the bag was very heavy''—the key words are underlined). The use of a female target speaker and a male masking speaker was likely to lead to less extensive informational masking than two same sex speakers (Brungart, 2001), but this was chosen to enable the instruction ''listen to the female speaker'' to be used throughout, rather than to train the subjects on the identity of the target speaker.

There were eight scans for each stimulus condition (noise masker and speech masker), with four different SNRs for each of these (+3, 0, −3, −6 dB for the speech-in-speech, +6, +3, 0, −3 dB for the speech-in-noise), presented twice in a random order. In the pretesting (see below) the stimuli were presented one target sentence (plus masker) at a time; during PET scanning the target sentences (plus maskers) ran continuously, for approximately 60 s. No overt response was required. During PET scanning, no sentences were repeated, either as targets or as maskers.

III. METHOD: BEHAVIORAL TESTING

Pilot testing was used to determine the intelligibility of the different masker and SNR conditions: ten normal-hearing adults (age range 26–50, five men, none of whom participated in the PET study) were presented with sentences over headphones and asked to repeat back the words that they could hear. This was done for a range of SNRs for both masker types (−12 to +9 dB SNR for the speech masker, −3-to +6 dB SNR for the noise masker). Sixteen sentences were used for each condition. These data were used to select the SNR conditions in which intelligibility was above a threshold (60%) to be used for the PET scanning.

The subjects for the PET study were presented with the stimuli prior to scanning. They were played individual BKB sentences and the masking stimuli over headphones and repeated back what they could hear. Eight sentences were presented per condition. Intelligibility was scored by an experimenter, who recorded the number of correct key words per condition. Since there was some variation in the number of key words per sentence, this score varied between a maximum of 18–20 key words per condition. This gave a score for each subject, masking condition, and SNR. The order of conditions was randomized.

IV. METHOD: PET SCANNING

Seven right-handed native English-speaking male volunteers were recruited and scanned. The mean age was 42, with a range of 35–52. Each participant gave informed consent prior to participation in the study, which was approved by the Research Ethics Committee of Imperial College School of Medicine/Hammersmith, Queen Charlotte's & Chelsea & Acton Hospitals. Permission to administer radioisotopes was given by the Department of Health (UK).

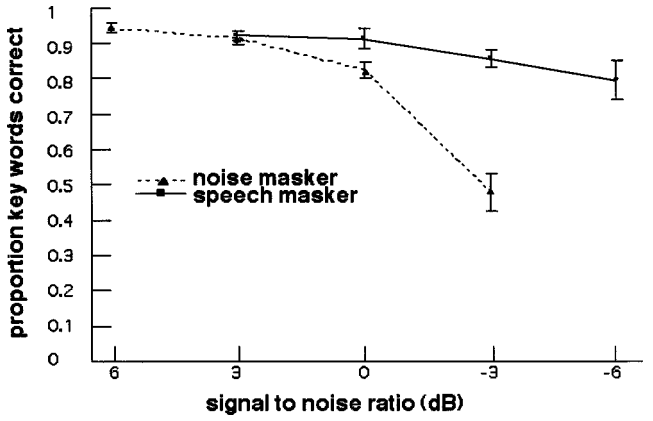


FIG. 1. The mean intelligibility of speech for the two maskers (speech and noise) as a function of SNR for the seven scanned subjects. Standard errors of the mean are shown by the error bars.

None of the subjects reported hearing problems and all were able to perceive speech in the different conditions during prescan training, though performance was poorer than the pilot subjects, such that performance in the lowest speech-in-noise SNR was lower than that in the pilot, with an average of 48% (see Fig. 1).

PET scanning was performed with a Siemens HR++ (966) PET scanner operated in high-sensitivity 3D mode. Sixteen scans were performed on each subject, using the oxygen-15-labeled water bolus technique. All subjects were scanned while lying supine in a darkened room with their eyes closed.

The stimuli were presented at a comfortable level determined for each subject, and this level was kept constant over the scanning sessions. The sentence presentations began 15 s before the scanning commenced, and each sentence presented was novel (i.e., there were no repeats). The subjects were instructed to listen passively to the female speaker ''for meaning'' in the scanning sessions. Passive listening reduces the likelihood that activation seen is due to controlled processing aspects of the task, which would be involved if the subjects were required to make explicit responses or try and remember the sentences they heard (Scott and Wise, 2003). Since this study is novel in focusing on normal speech perception in complex sound situations, this ensured that the activation seen was related to this, and not to some other aspect of the task requirements.

V. ANALYSIS

The images were analyzed using statistical parametric mapping (SPM99, Wellcome Department of Cognitive Neurology, http://www.fil.ion.ucl.ac.uk/spm), which allowed manipulation and statistical analysis of the grouped data. All scans from each subject were realigned to eliminate head movements between scans and normalized into a standard stereotactic space (the Montreal Neurological Institute template was used, which is constructed from anatomical MRI scans obtained on 305 normal subjects). Images were then

smoothed using an isotropic 10-mm, full width at half-maximum, Gaussian kernel, to allow for variation in gyral anatomy and to improve the signal-to-noise ratio.

In the statistical analysis of functional imaging data, the unit of analysis is the voxel, a cube of 2 mm^3. As voxel size is smaller than the resolution of the PET scanner, activity within adjacent voxels is not independent. Specific changes in regional cerebral blood flow (rCBF) were investigated by comparing the activity at each voxel in standardized space across scans (and, therefore, behavioral conditions). Specific, voxel-by-voxel analyses were performed using appropriate contrasts (e.g., which voxels are showing a greater response in speech masking contexts, relative to noise making contexts) to create statistical parametric maps of the t statistic, which were subsequently transformed into *Z* scores. The precise contrasts used are outlined in the next section. The threshold for significance was set at $P<0.05$, corrected for analyses across the whole volume of the brain ($P<0.000001$, uncorrected; *Z*-score >4.7). For contrasts where there is a prior hypothesis about the anatomical involvement, uncorrected p values of less than 0.0001 can be accepted.

The analysis included a blocked analysis of covariance (ANCOVA) with global counts as confound to remove the effect of global changes in perfusion across scans. This is necessary because there is considerable scan-to-scan variance in the number of counts (i.e., the precise amount of radio-labeled water infused) and in the blood flow of the subject (e.g., blood pressure tends to drop the longer one lies prone, and in a PET scan the subjects lie down for at least 1.5 h).

VI. RESULTS: BEHAVIORAL DATA

The intelligibility data from the pretesting of the scanned subjects are shown in Fig. 1. This demonstrates that, as predicted by previous studies, the speech-in-noise shows evidence of a level-dependent effect below a SNR of 3 dB; in contrast the speech-in-speech conditions show much less variation at 0 dB and below. A logistic regression was used to compare the trend in performance with SNR for the speech and noise masker. Analyses proceeded from the total number of key words correctly identified by each listener in each condition of SNR and masker type (six sentences containing a total of 18–20 key words), that is, 64 data points (8 listeners$\times$2 maskers ypes$\times$4 SNRs). SNR was the only continuous variate. The first analysis used listener, masker type, and SNR as predictors of performance (but with +3 dB added to the SNRs for the conditions with the speech masker, so that the 4 SNRs for each masker type were aligned). The triple interaction was not significant ($p\approx0.88$), but all the second-order interactions were ($p<0.005$). Significant terms involving listeners indicate that listeners varied in terms of their sensitivity to changes in SNR and masker type. More importantly, there was a significant interaction between masker type and SNR ($p<0.001$), indicating that performance depends upon SNR to a greater degree for the noise than the speech masker. A separate analysis of the role of SNR for the speech masker still found performance to depend significantly on SNR ($p<0.001$).

VII. RESULTS: PET SCANNING

Four different types of analyses were performed to investigate the neural correlates of the behavioural effects. The first were ''subtraction'' contrasts that reveal the activations which were greater when the participants listened to speech-in-speech, relative to speech-in-noise, and vice versa. These contrasts are insensitive to the different SNR levels. Since the design required the subjects to listen to the female speaker throughout the experiment, activations associated with the female speaker were ''subtracted'' out of the contrast. The second set of contrasts was parametric and investigated SNR-dependent responses by using the SNR as a covariate in each condition (speech-in-noise and speech-in-speech) separately. The third analysis was also parametric and used the intelligibility scores across all scans as a variable to identify neural activations that correlate positively with this. Finally, a statistical comparison was made between the speech-in-speech$>$speech-in-noise conditions and a previous study (Mummery *et al.*, 1999) which contrasted the perception of speech (single words) with a nonspeech baseline (signal correlated noise). This was an attempt to reveal the extent to which regions activated by ''unattended'' speech activated brain regions seen in passive speech percep-

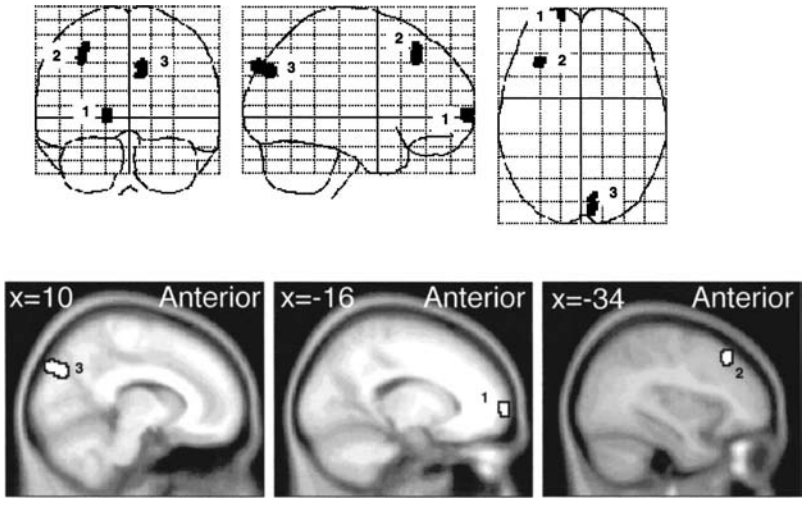


FIG. 2. Glass brain projections (upper panels) and sagittal slices on an average T1 weighted magnetic resonance imaging (MRI) image (lower panels) showing the activity seen for the speech-in-noise over speech-in-speech conditions, thresholded at $p<0.0001$, number of coactivated voxels >40. Labels: 1—left frontal pole, 2—left dorsolateral prefrontal cortex, 3—right posterior parietal cortex.

TABLE I. The locations, *Z* scores, corrected *p* values, and coordinates of the rCBF changes seen in the different contrasts (*indicates uncorrected *p* values).

Contrast	Region	*Z* score	*P* (corr)	*x*	*y*	*z*
Noise> speech	R posterior parietal	**5.22**	**0.009**	**10**	**−86**	**38**
	L rostral prefrontal cortex	**5.20**	**0.009**	**−16**	**68**	**2**
	L dorsolateral prefrontal cortex	**5.42**	**0.021**	**−34**	**28**	**44**
Speech> noise	L superior temporal gyrus	**7.30**	**0.000**	**−64**	**−20**	**2**
		6.73	0.000	−58	−8	2
		6.19	0.000	−68	−30	10
	R superior temporal gyrus	**7.18**	**0.000**	**64**	**−18**	**2**
		5.79	0.001	70	−26	4
		5.28	0.007	66	−8	0
Increasing intelligibility	L anterior superior temporal gyrus	**5.10**	**0.0015**	**−58**	**0**	**0**
(speech>SCN) +(sp-in-speech>sp-in-noise)	L lateral STG	**5.38**	**0.022**	**−52**	**−14**	**2**
	L posterior STS	5.23	0.039	−56	−32	6
	L lateral STG	5.16	0.053	−66	−16	0
	R lateral STG	4.99	0.096	66	−24	2
	L anterior STG	4.90	0.000*	−60	8	−2

tion. An important proviso is that these two studies were performed on different PET scanners, which may make the comparison less sensitive.

A comparison of the activity seen in the speech-in-noise condition, relative to the speech-in-speech condition, revealed SNR-independent responses in the left frontal pole, left dorsolateral prefrontal cortex, and the right posterior parietal cortex. (Fig. 2, Table I) In contrast, strong responses were seen for speech-in-speech over speech-in-noise in the left and right lateral superior temporal gyri (STG) and sulci (STS), running posterior, lateral and anterior to primary auditory cortex (Fig. 3, Table I). The activation also extends into Heschl's gyrus on the right, although this is not a separate peak of activation. Since the behavioral data indicates

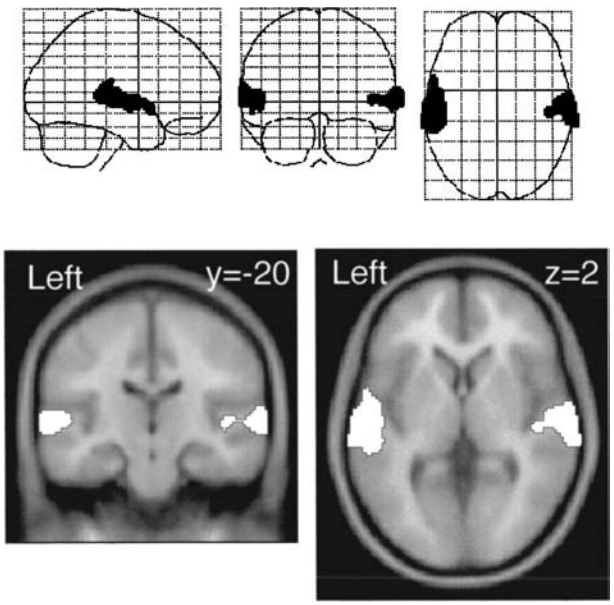


FIG. 3. Neural activity in the speech-in-speech condition, relative to the speech-in-noise condition; the rCBF changes are shown on glass brain projections and coronal and transaxial slices of an average T1 weighted MRI image, thresholded at $p<0.0001$, number of co-activated voxels >40.

that the lowest SNR for speech-in-noise (−3 dB) results in the poorest comprehension (see Fig. 1), the analysis was repeated excluding this condition and thus avoiding any activations due to gross comprehension differences between the speech-in-noise and in the speech-in-speech conditions. With this condition excluded, the peak changes little, moving just 2 mm up in the *z* plane, with the other coordinates remaining unaffected ($Z=7.04$).

The analysis of rCBF changes with SNR revealed no significant changes for speech-in-speech, either for SNR increases or decreases. SNR-dependent analysis of speech-in-noise revealed activations in left inferior prefrontal cortex [−42, 20, −12, $Z=4.15$, $P(uncorrected)<0.001$] and left dorso-medial premotor area [−14, 2, 70, $Z=4.15$, $P(uncorrected)<0.001$] (Fig. 4). The rCBF levels appear to vary linearly with the decreasing SNR conditions in both regions [Figs. 4(a) and (b)], and reflect neural responses that increase as the listening task becomes harder. These results are included, however, with the proviso that the *p* values fall below the level of significance for whole brain comparisons. There was no significant rCBF change associated with increasing SNR levels (i.e., as speech perception becomes easier).

The use of the subjects' average intelligibility scores across all the masker and SNR conditions as a covariate revealed a left lateralized peak, in anterior STG (Fig. 5, Table I), where activity increased positively with intelligibility. To demonstrate that this effect was seen across all the subjects the analysis was repeated with the subjects' individual intelligibility scores entered as subject specific covariates, which gave one peak at the same location in the anterior STG (−58, 0, 0, $Z=4.59$). Individual subjects' rCBF values for this peak voxel were plotted against their intelligibility scores and the regression lines for each plot. This showed that for all but one of the subjects the relationship is a positive one, which is why a positive correlation between rCBF and intelligibility comes out in the mean intelligibility analysis. The R^2 values (and corresponding *p* values) for the linear

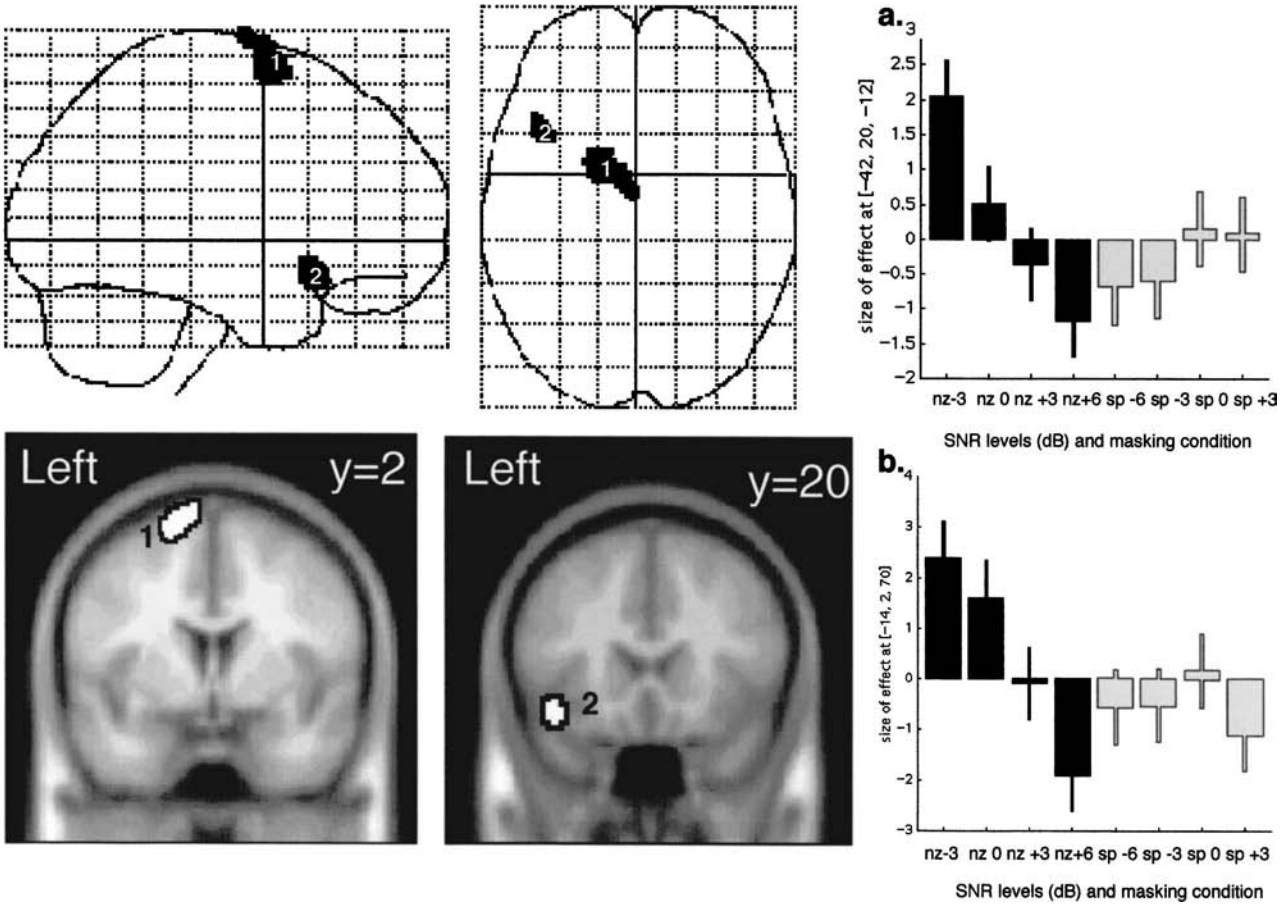


FIG. 4. Glass brain projections (upper left panels) and coronal slices of an average T1 weighted MRI image of rCBF changes that correlate with the difficulty of the speech-in-noise condition—i.e., negatively correlated with the SNR values, thresholded at $p<0.001$, number of coactivated voxels >40. Key: 1—left dorso-medial premotor cortex, 2—left inferior lateral prefrontal cortex. The graphs on the right show the rCBF changes associated with the different SNR values for the speech-in-noise condition at the peak voxel in the inferior prefrontal region (a) and the dorso-medial premotor activation (b).

regression for each subject (1–7) respectively are 0.123 ($p=0.18$), 0.447 ($p=0.0042$), 0.252 ($p=0.05$), 0.436 ($p=0.005$), 0.074 ($p=0.30$), 0.466 ($p=0.003$), 0.0014 ($p=0.87$). These R^2 values indicate there is a considerable amount of rCBF variation accounted for by the intelligibility of the sentences, for five out of the seven subjects, and that this was significant at $p<0.05$ for four of the seven subjects.

A second level random effects analysis was used to compare the activation for speech in speech>speech-in-noise and the speech>signal correlated noise contrast from Mummery *et al.* (1999) (also smoothed to 10 mm). This tested for regions coactivated by both contrasts, and for those significantly more activated by the speech in speech >speech-in-noise contrast. This revealed extensive coactivations in bilateral STG/STS and the supratemporal planes, with peaks in left posterior STS, bilateral STG, and left anterior STS. (Fig. 6, Table I)

VIII. DISCUSSION

The results of this study reveal a clear difference between the neural processing of speech when it is masked by speech versus noise. The former is associated with extensive activation in bilateral superior temporal gyri, the latter with the recruitment of brain regions remote from those classically associated with speech perception. Previous behavioral studies have suggested that speech and noise act as maskers in distinctly different ways (e.g., Brungart, 2001), and this neuroimaging evidence is consistent with such observations.

The use of unmodulated speech-spectrum noise as a masking stimulus, although at SNR levels above those that eliminate comprehension, shows a distributed network of neural regions, consisting of right parietal cortex and left prefrontal cortex (when contrasted with listening to speech-in-speech). The responses in these regions are independent of SNR level. Such patterns of activation are not seen in normal speech perception when a passive listening task is used (e.g., Binder *et al.*, 2000; Mummery *et al.*, 1999; Scott *et al.*, 2000, Wise *et al.*, 1991; 2001); this suggests that these areas are recruited to facilitate the perception of speech specifically in the context of unmodulated noise, which makes the speech difficult to hear due to masking at the auditory periphery. In this context the activation might be associated with some degree of attention; indeed prefrontal and parietal activations are associated with online, controlled cognitive processing and the involvement of explicit attentional mechanisms, albeit those that are not specific to auditory processing [e.g., prospective memory (Burgess *et al.*, 2001); auditory vigilance (Paus *et al.*, 1997); rapid visual information processing (Coull *et al.*, 1996)].

When the subjects listen to the female speaker in the

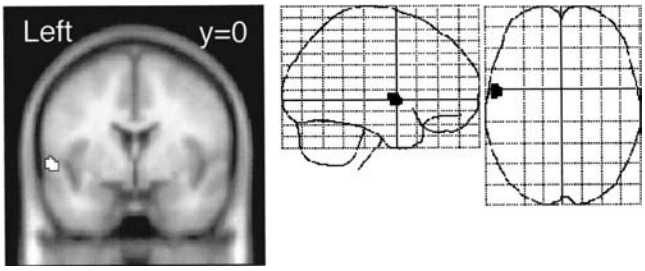


FIG. 5. Coronal slice of an average T1 weighted MRI image and glass brain projections of the rCBF regions that correlate with the increases in intelligibility, across all conditions, thresholded at $p<0.0001$ with number of coactivated voxels >40.

context of another (male) speaker a different pattern of activation is seen. Despite the fact that this contrast, by comparing the perception of speech-in-speech with speech in noise, ''subtracts'' the mental process of perceiving speech (as it is present in both conditions), the activation in left and right STG/STS is extensive, and independent of the SNR level (Fig. 3). There is considerable correspondence between the glass brain images for this contrast and that of a PET study contrasting speech with signal correlated noise (SCN); the activations extend along the lateral STG/STS, and their extents are similar in the anterior and posterior dimensions. This similarity was confirmed by a statistical comparison with this previous PET study (Fig. 6). Regions which are activated by both speech>SCN and speech-in-speech

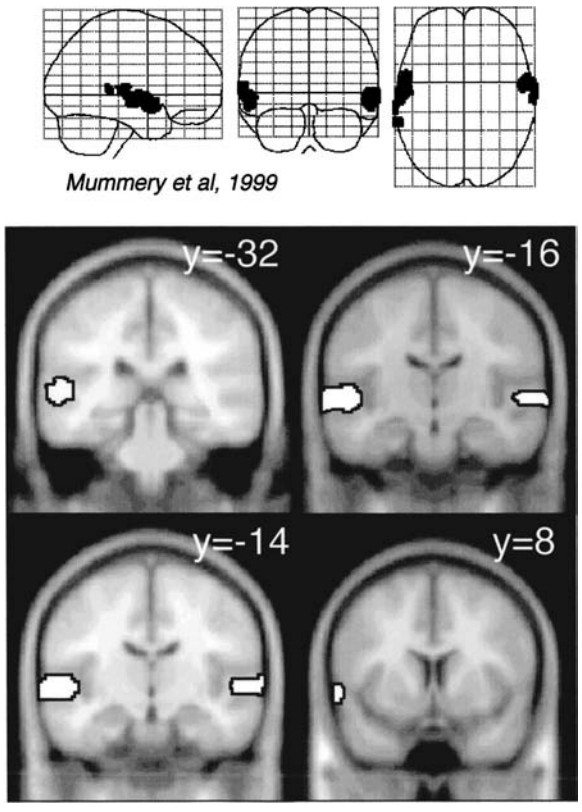


FIG. 6. Glass brain projections of the speech specific responses seen in an earlier study [reanalysis of Mummery *et al.* (1999)] are shown on the top panels: The activations seen for speech-in-speech over speech-in-noise and for speech>signal correlated noise are shown in the lower figures, on coronal slices of an average T1 weighted MRI image. Both sets of activations are thresholded at $p<0.0001$ with number of co-activated voxels >40.

>speech-in-noise run bilaterally along the supra temporal plane and lateral STG/STS, with greater anterior and posterior extent on the left.

There are several potential explanations of this result. The peaks in posterior STG are consistent with claims that posterior auditory fields are involved in the analysis of auditory objects and their location (Griffiths and Warren, 2002; Zatorre *et al.*, 2002), even though there were no cues to spatial location in this study. This finding may thus reflect the neural basis of grouping auditory objects, here different speakers, a processing demand which is increased in the speech-in-speech condition. The peaks in bilateral lateral STG/STS and left anterior STG are consistent with the ''unattended'' speech being processed along the same stream of processing as attended speech. This argument suggests regions antero-lateral to PAC may be involved in processing more than one concurrent speech source, and that informational masking occurs as a result of these competing speech related cues. The use of an informational masking signal which is not intelligible, but which is acoustically as complex in its structure as speech [e.g., spectrally rotated speech (Scott *et al.*, 2000)] will be able to reveal the extent to which such activation is a result of the acoustic overlap between the two signals, or the semantic content of the masking speech.

A different interpretation is that ''glimpsing'' of the target signal, due to amplitude modulations in the masking speech, may lead to greater activation in these regions than is seen when perceiving speech in unmodulated noise. This would result in increased activations of speech processing regions as more of the speech signal is available for cortical processing (due to the ''gaps'' in the masker). This can be explicitly addressed in further studies using amplitude-modulated noise maskers (to enable some ''glimpsing''). It is also the case that the speech-in-speech condition is considerably more complex than speech alone, in terms of its modulation spectrum and spectral profile: future studies with unintelligible masking stimuli as complex as speech will be able to address the role of the acoustic profile of the masking stimulus in the pattern of activation seen.

There is some evidence from this study for an involvement of primary auditory cortex in informational or energetic masking: the speech-in-speech>speech-in noise contrast shows right STG activation that extends medially into Heschl's gyrus, which is the location of primary auditory cortex in man (Fig. 3). However, this was not a separate peak of activation, making it harder to be certain about the validity of this finding. This equivocal result could, however, be due to the power of the analysis, as a consequence of resolution of the technique and the design of the current study. Blood flow measures may not be able to resolve small changes due to such modulation in a study where auditory stimulation is present in each condition.

One tentative conclusion, therefore, is that in the informational masking conditions there is some evidence that both auditory object segregation regions and speech processing regions are involved, and that these may contribute to the central auditory processes associated with informational masking. This may also suggest that the bilateral STG/STS regions commonly seen in functional imaging studies of

speech perception are also capable of processing in parallel other, unattended, speech information; the converse of this is that when the female speaker is ''streamed'' out of the auditory scene, the selection is not occurring ''early'' in the processing of the perceptual stream. The activations common to speech perception and speech-in-speech>speech-in-noise run along the STG/STS into regions which, on the left, are associated with the processing of intelligible speech (Scott *et al.*, 2000), consistent with the suggestion that the unattended voice is processed, to some degree, for meaning. Further studies will be able to address how these posterior and lateral auditory regions interact in informational masking.

Several cognitive processing tasks have revealed the consequences of obligatory perceptual processing of unattended speech. Early studies indicated that selective attention can be directed to the meaning of speech, with the potential for interference if concurrent ''unattended'' speech overlaps in semantic content (e.g., Treisman, 1960). There is also evidence that that concurrent irrelevant speech can disrupt performance on verbal working memory tasks (Tremblay *et al.*, 2000). This suggests that the perceptual competition underlying informational masking can also affect the cognitive processes dependent upon speech perception. There is also evidence that intelligible irrelevant speech is processed for meaning: recent work, for example, has shown that the meaning of unattended speech can interfere with semantic processing (e.g., Neely and LeCompte, 1999). The current findings may demonstrate a candidate neural basis for these behavioral effects. Importantly, as with informational masking, aspects of irrelevant speech disruptions in working memory tasks appear not to be speech specific, and effects can be seen with tone sweeps and reversed speech, although there is a relationship between the acoustic features of the irrelevant signal and the amount of distraction it causes (e.g., Tremblay *et al.*, 2000). As noted, further functional imaging studies can determine the extent to which the to-be-ignored voice is processed because it is meaningful speech, or because it is acoustically similar to speech.

The investigation of a SNR-dependent response for speech-in-speech revealed no activation that correlated with the SNR, although this is arguing to a null result and future work with more sensitive techniques [e.g., same sex speakers, which will increase the amount of informational masking (Brungart, 2001)] may find a difference. This finding is, however, consistent with the hypothesis that there would be no such response, since the behavioral data indicates that informational masking is SNR independent at SNRs of 0 dB and below (Brungart, 2001). In contrast, there were activations in the speech-in-noise that correlated with SNR, consistent with the original hypothesis and the behavioral data. Since there was no prediction about likely regions associated with SNR-dependent responses, and the activations did not survive correction for whole brain comparisons, the activations seen in left dorso-medial premotor cortex and left inferior prefrontal cortex (shown in Fig. 4) must be treated with some caution. However, there is evidence for an involvement of both these regions in speech processing. In a recent paper on speech production (Blank *et al.*, 2002) extensive medial prefrontal activation was associated with the production of ''propositional'' speech, (e.g., a verbal description of a relative that you see often but don't live with), in contrast to ''automatic'' speech (e.g., speaking a very familiar nursery rhyme repeatedly). This activation extended dorsally to the premotor region which shows the SNR-dependent response; as the speech is harder to hear in the noise, there is greater activation in this region. This premotor response potentially reflects the use of articulatory strategies (i.e., ''sounding out'' the heard words), which may or may not be explicit (i.e., associated with a deliberate strategy), and which are recruited to facilitate speech perception in the context of energetic masking. Previous studies have implicated ''mirror'' neurones (Rizzolatti and Arbib, 1998) in lateral premotor regions (posterior Broca's area) in aspects of speech perception when the task makes more motoric demands, e.g., segmentation (Burton *et al.*, 2000). The activation seen in this study is dorsal and medial to such activations, and this possibly reflects the fact that subjects simply had to comprehend the speech in the current study without doing such complex phoneme monitoring. In other words, no meta-linguistic processing was required.

With respect to the inferior prefrontal cortical response, another recent study (Crinion *et al.*, 2003) showed a response in ventral prefrontal cortex to hearing speech (children's stories) compared to reversed speech. The peak was very close (−44 26 −16) to that seen here. Crinion *et al.* (2003) associated this response with prefrontal regions that receive projections from anterior temporal lobe regions implicated in the perceptual processing of intelligible speech (Scott *et al.*, 2000) and they identified this region as a more executive component of story comprehension. Certainly the neuroanatomy is consistent with this claim, and it is intriguing to speculate that the ventral prefrontal activation seen in the current study demonstrates ''top down'' efforts to support speech comprehension with semantic information.

The correlation of intelligibility scores with rCBF across both conditions revealed a left lateralized response, lying lateral and anterior to primary auditory cortex in the STG. A previous study from our group, which investigated the neural correlates of speech processing while controlling for auditory complexity, showed responses associated with intelligibility of speech in left anterior STS (Scott *et al.*, 2000). The peak is superior and posterior to the most anterior peak shown in the Scott *et al.* (2000) study, which may be due to the task; subjects in the previous study were not forced to stream the speech out from a noisy background. Another difference is one of sensitivity, since more of the variation behaviorally comes from reduced comprehension in the most difficult speech-in-noise condition, and the other conditions do not differ greatly. Nonetheless, the activation which is associated with increasing comprehension of the attended speech lies in the anterior temporal lobe and not in posterior temporal lobe regions often claimed to be central for the processing of speech for meaning (Hickok and Poeppel, 2000). In the context of the current study and our work on intelligibility, this confirms that anterior temporal lobe regions are associated with intelligibility in speech.

This study thus presents evidence that different masking contexts for speech perception recruit different neural sys-

tems. Regions in rostral and dorsolateral prefrontal cortex and posterior parietal cortex are recruited, in a SNR-independent fashion, when subjects listen to speech-in-noise. In contrast, speech-in-speech activates bilateral STG/STS, in addition to the activation associated with the perception of the attended speech, indicating parallel processing of the speakers. Further studies are needed to determine whether this occurs because the unattended speech is meaningful, or whether it is due to more basic acoustic properties of the masking signal. Future studies will also be able to determine whether the SNR-dependent effects seen for speech-in-noise reflect the automatic recruitment of articulatory and semantic mechanisms in difficult speech perceptual conditions, or whether these reflect more explicit and deliberate cognitive strategies. The overall correlation of intelligibility of the signal with the activity in left anterior STG is further evidence for the role of anterior regions in the comprehension of intelligible speech.

ACKNOWLEDGMENTS

SKS and RJSW are both funded by grants from the Wellcome Trust. We would like to thank Phil Beaman for helpful comments on the manuscript.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., and Possing, E. T. (**2000**). ''Human temporal lobe activation by speech and nonspeech sounds,'' Cereb. Cortex **10**, 512–528.

Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., and Wise, R. J. S. (**2002**). ''Propositional speech production: Broca, Wernicke and beyond,'' Brain **125**, 1829–1838.

Brungart, D. S. (**2001**). ''Informational and energetic masking effect in the perception of two simultaneous talkers,'' J. Acoust. Soc. Am. **109**, 1101–1109.

Brungart, D. S., and Simpson, B. D. (**2002**). ''Within-ear and across-ear interference in a cocktail-party listening task,'' J. Acoust. Soc. Am. **112**, 2985–2995.

Burgess, P. W., Quayle, A., and Frith, C. D. (**2001**). ''Brain regions involved in prospective memory as determiend by positron emission tomography,'' Neuropsychologia **39**, 545–555.

Burton, M. W., Small, S. L., and Blumstein, S. E. (**2000**). ''The role of segmentation in phonological processing: An fMRI investigation,'' J. Cogn Neurosci. **12**, 679–690.

Cherry, E. C. (**1953**). ''Some experiments on the recognition of speech, with one or two ears,'' J. Acoust. Soc. Am. **25**, 975–979.

Conway, A. R., Cowan, N., and Bunting, M. F. (**2001**). ''The cocktail party phenomenon revisited: the importance of working memory capacity,'' Psychonomic Bull. Rev. **8**, 331–335.

Coull, J. T., Frith, C. D., Frackowiak, R. S. J., and Grasby, P. M. (**1996**). ''A fronto-parietal network for rapid visual information processing: A PET study of sustained attention and working memory,'' Neuropsychologia **34**, 1085–1095.

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., Wise, R. J. (**2003**). ''Temporal lobe regions engaged during normal speech comprehension.'' Brain, **126**, 1193–1201.

Davis, M. H., and Johnsrude, I. S. (**2003**). ''Hierarchical processing in spoken language comprehension,'' J. Neurosci. **23**, 3423–3431.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., Matthews, P. M., and Tyler, L. K. (**2000**). ''Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task,'' Neuroimage **11**, 589–600.

Dirks, D., and Bower, D. (**1968**). ''Masking effects of speech competing messages,'' J. Speech Hear. Res. **12**, 229–245.

Festen, J., and Plomp, R. (**1990**). ''Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing,'' J. Acoust. Soc. Am. **88**, 1725–1736.

Foster, J. R., Summerfield, A. Q., Marshall, D. H., Palmer, L., Ball, V., and Rosen, S. (**1993**). ''Lip-Reading the BKB sentence lists—corrections for list and practice effects,'' Br. J. Audiol. **27**, 233–246.

French, N., and Steinberg, J. (**1947**). ''Factors governing the intelligibility of speech sounds,'' J. Acoust. Soc. Am. **19**, 90–119.

Griffiths, T. D., and Warren, J. D. (**2002**). ''The planum temporale as a computational hub,'' Trends Neurosci. **25**, 348–353.

Hickok, G., and Poeppel, D. (**2000**). ''Towards a functional neuroanatomy of speech perception,'' Trends Cogn. Sci. **4**, 131–138.

MacLeod, A., and Summerfield, Q. (**1987**). ''Quantifying the contribution of vision to speech perception in noise,'' Br. J. Audiol. **21**, 131–141.

Moray, N. (**1959**). ''Attention in dichotic listening: Affective cues and the influence of instructions,'' Q. J. Exp. Psychol. **11**, 56–60.

Mummery, C. J., Ashburner, J., Scott, S. K., and Wise, R. J. S. (**1999**). ''Functional neuroimaging of speech perception in six normal and two aphasic patients,'' J. Acoust. Soc. Am. **106**, 449–457.

Neely, C. B., and LeCompte, D. C. (**1999**). ''The importance of semantic similarity to the irrelevant speech effect,'' Mem. Cognit. **27**, 37–44.

Paus, T., Zatorre, R. J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., and Evans, A. C. (**1997**). ''Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task,'' J. Cogn Neurosci. **9**, 392–408.

Rizzolatti, G., and Arbib, M. A. (**1998**). ''Language within our grasp,'' Trends Neurosci. **21**, 188–194.

Schroeder, M. R. (**1969**). ''Reference signal for signal quality studies,'' J. Acoust. Soc. Am. **44**, 1735–1736.

Scott, S. K., and Johnsrude, I. S. (**2003**). ''The neuroanatomical and functional organization of speech perception,'' Trends Neurosci. **26**, 100–107.

Scott, S. K., and Wise, R. J. S. (**in press**). ''Functional Imaging and Language: A Critical Guide to Methodology and Analysis,'' to appear in Speech Commun.

Scott, S. K., Blank, S. C., Rosen, S., and Wise, R. J. S. (**2000**). ''Identification of a pathway for intelligible speech in the left temporal lobe,'' Brain **123**, 2400–2406.

Treisman, A. M. (**1960**). ''Contextual cues in selective listening,'' Q. J. Exp. Psychol. **12**, 242–248.

Tremblay, S., Nicholls, A. P., Alford, D., and Jones, D. M. (**2000**). ''The irrelevant sound effect: does speech play a special role?'' J. Exp. Psychol. Learn. Mem. Cogn. **26**, 1750–1754.

Ulanovsky, N., Las, L., and Nelken, I. (**2003**). ''Processing of low-probability sounds by cortical neurons,'' Nat. Neurosci. **6**, 391–398.

Wise, R. J. S., Scott, S. K., Blank, S. C., Mummery, C. J., and Warburton, E. (**2001**). ''Identifying separate neural sub-systems within, ''Wernicke's area,'' Brain **124**, 83–95.

Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., and Frackowiak, R. (**1991**). ''Distribution of cortical neural networks involved in word comprehension and word retrieval,'' Brain **114**, 1803–1817.

Zatorre, R. J., Bouffard, M., Ahad, P., and Belin, P. (**2002**). ''Where is 'where' in the human auditory cortex?'' Nat. Neurosci. **5**, 905–909.

A molecular description of profile analysis: Decision weights and internal noise[a)]

Bruce G. Berg[b)]
Department of Cognitive Science, University of California, Irvine, California 92697

(Received 5 July 2002; revised 29 October 2003; accepted 17 November 2003)

Systematic inefficiencies and internal noise in a spectral profile discrimination task were investigated. Listeners detected a 1000-Hz sinusoid added in-phase to the central component of a complex consisting of 11 equal-intensity sinusoids. Parameters for a channel model that employs decision weights and internal noise were estimated with molecular psychophysical techniques. Maximum likelihood predictions of the model were generally within a few decibels of observed thresholds. The degree to which an assumption of ideal weights leads to overestimation of internal noise was also assessed. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639904]

PACS numbers: 43.66.Ba, 43.66.Fe [MRL] Pages: 822–829

I. INTRODUCTION

Swets *et al.* (1959) proposed two sources of inefficiency in the detection of multiple signals in noise—the inappropriate integration of information from multiple sources (e.g., frequency channels) and the inherent biological variability of the listener (i.e., internal noise). These principles can also be applied to spectral discrimination tasks. Techniques have been developed to estimate internal noise (Swets *et al.*, 1959; Watson, 1962; Green, 1964; Spiegal and Green, 1981; Robinson and Berg, 1986; Berg, 1987; Lutfi, 1989; Dai, 1994; Lentz and Richards, 1997; Murray *et al.*, 2002) and to investigate the implicit decision strategies of listeners (Ahumada, 1967; Gilkey and Robinson, 1986; Berg, 1989; Richards and Zhu, 1994; Lutfi, 1995; Murray *et al.*, 2002). That the techniques have been developed more or less independently is not without consequence. Estimates of internal noise are typically obtained under the assumption that the decision strategy of the listener is ideal. This might hold for simple, single-channel stimuli, but is unlikely to hold for complex sounds which require central integration of information from different channels or different time segments. Listeners almost always exhibit some degree of inefficiency in their integration strategies, as shown in tone-in-noise detection (e.g., Ahumada and Lovell, 1971; Ahumada *et al.*, 1975; Richards and Buss, 1996), sample discrimination (Sorkin *et al.*, 1987; Lutfi, 1989, 1992a, 1995), spectral profile discrimination (Berg and Green, 1990, 1991; Green and Berg, 1991; Dai and Berg, 1992), intensity discrimination (Doherty and Lutfi, 1996, 1999), discrimination of narrow-band spectra (Berg *et al.*, 1992; Southworth and Berg, 1995), temporal pattern discrimination (Sorkin, 1990; Sadralodabai and Sorkin, 1999), and discrimination of multiple element visual displays (Sorkin *et al.*, 1991; Montgomery, 1991). A possible exception is Lentz and Richards (1997), in which data from a number of profile discrimination conditions were fit quite well with a single parameter representing internal noise.

One objective of this paper is to quantify and correct the overestimation of internal noise whenever an optimal decision strategy is assumed for a nonoptimal observer. Although arguments are developed within the theoretical framework of spectral profile discrimination (Green, 1988), they apply to any discrimination task that can be broadly construed as entailing multiple observations. Tanner and Birdsall (1958) were among the first to consider internal noise in the context of the theory of signal detection (see Green and Swets, 1966). In standard notation, if ideal sensitivity is $d' = (\mu_s - \mu_n)/\sigma$, then listener sensitivity can be expressed as $d'_{obs} = (\mu_s - \mu_n)/(\sigma^2 + \sigma_I^2)^{1/2}$, where σ_I^2 represents internal noise. This leads to the efficiency measure

$$\eta = \frac{(d'_{obs})^2}{(d')^2} = \frac{\sigma^2}{\sigma^2 + \sigma_I^2}. \tag{1}$$

By this conceptualization, any loss in performance relative to the ideal is due entirely to the inherent variability of the listener. Sources of variability can be quite diverse, such as a pliable memory representation of an expected signal, momentary fluctuations in the biochemical state of auditory fibers, or even an inadvertent scuffling of a chair. Green (1964) described internal noise as any event which entails a random component.

A conceptualization more consonant with the ideas of Swets *et al.* (1959) is favored here, one for which efficiency is a function of both internal noise and the decision strategy of the listener. Green (1964) made the distinction between random and systematic inefficiencies, and characterized sequential response dependencies as one type of systematic inefficiency (e.g., Treisman and Williams, 1984). The concern here is with the systematic inefficiency in combining information across frequency channels. Disregarding a channel that transmits relevant information is a systematic error, as is reliance on an uninformative channel.

It is often the case that the ideal decision variable (DV) in multiple observation tasks is a weighted sum of the observations, $\mathrm{DV} = \Sigma a_j x_j$, where x_j is the jth observation and a_j is the associated decision weight. By definition, a set of weights represents a listener's decision strategy. Letting d'_W represent the sensitivity of a hypothetical observer that uses

[a)]Presented in part at the 1992 Fall Meeting of the Acoustical Society of America [J. Acoust. Soc. Am. **92**, 2316 (1992)].
[b)]Electronic mail: bgberg@uci.edu

the estimated weights of a listener, but that otherwise exhibits no inefficiency, Berg (1990) defined weighting efficiency as $\eta_W=(d'_W/d')^2$, a measure representing the loss of efficiency attributable to nonoptimal weights. A second measure, $\eta_N=(d'_{obs}/d'_W)^2$, represents the loss of efficiency attributable to internal noise. Overall efficiency can thus be partitioned into random and systematic components,

$$\eta=\eta_N\eta_W=\frac{(d'_{obs})^2}{(d'_W)^2}\frac{(d'_W)^2}{(d')^2}=\frac{(d'_{obs})^2}{(d')^2}. \quad (2)$$

In a profile discrimination task, Berg and Green (1991) found that η_W decreased as the complexity of spectral changes increased. Another common finding is that listeners with the lowest thresholds exhibit the greatest η_W (Berg and Green, 1990, 1991; Green and Berg, 1991). As an explanation for the finding that signals added at either edge of a spectral profile are more difficult to detect (Green and Mason, 1985), Bernstein and Green (1987) postulated that internal noise is greater for high and low frequency channels than for mid frequency channels [see Watson *et al.* (1972) for a related discussion]. Green and Berg (1991), on the other hand, presented evidence that the ''bowl-effect'' stems from a decrease in η_W when the signal is added near a spectral edge, though it may be the case that both factors contribute to lower performance. To summarize, these studies established that idiosyncratic decision strategies have a measurable influence on performance in profile discrimination tasks.

A method for estimating internal noise (Berg, 1987; Sorkin *et al.*, 1987) and a method for estimating decision weights (Berg, 1989; Berg and Green, 1990) are consolidated here. The technique combines what Green (1964) has termed molar and molecular psychophysics. Parameters are estimated at a molecular level and then used to make predictions at a molar level. A straightforward delineation of the two levels of analysis is that molar psychophysical measures, such as $P(c)$ or d', consider the accuracy of responses, whereas molecular psychophysical measures consider the stimulus-response relationship with no regard to ''correctness.'' Green, for instance, estimated the percent-of-agreement across two exact replications of a block of stimuli, using only those trials consisting of two noise bursts, neither of which contained a signal. Gilkey and Robinson (1986) fit psychometric functions with the ordinate defined as the probability of a ''yes'' response, $p(y)$, rather than percent correct. Similarly, Berg's (1989) COSS analysis[1] is based on $p(y)$ rather than $P(c)$.

Molecular psychophysical techniques provide descriptions of listening behavior that are unattainable with molar psychophysics. Individuals often attend to very different aspects of the available acoustic information, yet show similar thresholds (e.g., Berg and Green, 1991). Molecular techniques also allow models to be evaluated on more than the accuracy of their threshold predictions. It is shown below that a model with a single parameter representing internal noise is as accurate at predicting thresholds as a model that incorporates both internal noise and a listener's decision weights. The latter, however, is more consistent with what may be termed the microstructure of the data (e.g., the slope of a COSS function).

II. METHOD

A channel model (Durlach *et al.* 1986) was used to describe performance in a profile discrimination task. The model was modified in order to incorporate estimates of internal noise and decision weights. The experiment is described first, followed by descriptions of the model, and a summary of data analysis techniques. Although derivations of the quantitative techniques have appeared elsewhere, this is the first collective treatment.

A. Stimuli

The standard consisted of 11 sinusoids, each with intensity $L+x_j$, where L is the nominal intensity and x_j is a small, in-phase level perturbation. Weight estimates are based on the effects of the perturbations. For each sinusoid on each stimulus presentation, x_j was randomly sampled from a normal distribution (on a decibel scale) with a mean of zero and a variance of 1 dB. L was the same for all sinusoids and was sampled on every presentation from a uniform distribution with a range of 55 to 75 dB SPL. Frequencies ranged from 200 Hz to 5 kHz and were equally spaced to the nearest 10 Hz on a log-frequency scale. The phase of each component was randomly sampled from a uniform distribution. The signal was a 1000-Hz sinusoid added in-phase to the 1000-Hz component of the standard.

A trial consisted of two 500-ms intervals separated by 500 ms. The signal was added to the standard in either the first or second interval with equal probability. Response time was unlimited and visual feedback was presented after each trial. Subjects used a keyboard to report their decisions about which interval contained the signal.

Digitally produced waveforms were shaped by 20-ms cosine-squared onset/offset ramps and played through a digital-to-analog converter at 25 kHz; the anti-aliasing filter had a cutoff of 6.5 kHz. Manual attenuators were used for calibration and programmable attenuators (TDT PA4) controlled the level of the standard and signal. Sounds were delivered diotically over Sennheiser HD414SL headphones to a listener seated in a single-walled, IAC chamber.

B. Procedure

Data were collected in two stages. The first was used to obtain a fixed signal level that was subsequently needed for data collection during the second stage. All reported data were from the second stage.

For the first stage, thresholds were estimated with a 2IFC, staircase procedure (Levitt, 1971). Starting approximately 10 dB above threshold, the signal level was reduced by 4 dB after two consecutive correct responses and increased by 4 dB after each incorrect response. After the third reversal, the stepsize was reduced to 2 dB. A threshold was estimated by averaging the signal levels at all reversals of the adaptive track up to the last even-numbered reversal, excluding the first three. The mean threshold across ten 50-trial blocks determined the fixed signal level for the second stage.

A procedural modification was needed for estimating weights because the analysis requires a fixed signal level. Although the most statistically efficient signal level has yet to be derived, previous investigations have shown success by fixing the signal at a level yielding 71% correct, as obtained in the first stage. Another modification derives from undocumented observations that the performance of some listeners deteriorates when blocks consist entirely of trials with the signal level fixed at threshold. Interleaving fixed and adaptive trials appears to circumvent this problem, perhaps because memory is refreshed by signal levels at the peaks of the staircase. Blocks were increased to 100 trials for the second stage. The first five trials of each block were adaptive. Thereafter, adaptive and fixed trials were selected with equal probability. Signal levels for the adaptive track followed the schedule described above. Weights were estimated from fixed trials and thresholds were estimated from adaptive trials. Additional fixed trials were gained by setting the initial level of the adaptive track to a value that would lead to a ''step'' at the fixed level. One drawback of the analysis is the amount of data required. Each listener completed at least 42 blocks, with a mode of 48 blocks. Reported thresholds are averaged across all blocks of the second stage.

C. Subjects

Five undergraduate students at the University of Florida, all younger than 25 years old, and the author participated as listeners. All had normal pure-tone thresholds in the frequency range of the stimulus. Students received monetary compensation and were trained on the task for a minimum of 4 h prior to the experiment. The author's data were previously reported in Berg and Green (1990), though without application of the model.

D. Data analysis

1. Molar description

The channel model provides the theoretical framework for fitting threshold data. Quasi-independence between channels is assumed, in the sense that the separation between spectral components is large enough so that none fall within the same critical band. The roving-level procedure, however, introduces strong correlations among the channels. For the original model, the covariances between channels are factors in a multivariate decision variable known as Mahalonobis distance (Mahalanobis, 1936; c.f. Arnold, 1981). Berg and Green (1990) showed that the DV can be expressed as a weighted sum of channel outputs in a profile discrimination task, such that sensitivity can be expressed as

$$d'_{obs}=\frac{\sqrt{2}\Sigma_{j=1}^{n}a_j\Delta L_j}{\sqrt{\Sigma_{j=1}^{n}a_j^2\sigma_x^2+a_L^2\sigma_L^2+\sigma_I^2}}. \quad (3)$$

Stimulus uncertainty is a function of the variance of the perturbations, $\sigma_x^2=1$ dB, and the variance of the roving level, $\sigma_L^2=33$ dB. Estimated parameters are the decision weights, a_j, $j=1,...,n$, internal noise, σ_I^2, and the level weight, a_L, which is thought to quantify the relative influence of the absolute level. In theory, $a_L=\Sigma a_j$, although in practice, a_L is estimated directly, as discussed below.

The predicted variable is ΔL_j, representing the signal level of the jth sinusoid at threshold. For the current task, $\Delta L_j=0$ for $j\neq 6$. The subscript allows generalizations to signals of greater complexity, such as the sinusoidal-shaped spectra used by Berg and Green (1991). For the calculations, ΔL_j is expressed as the level difference, $20\log(1+\Delta p/p)$. By convention, however, reported thresholds are converted to the signal-to-standard ratio, $20\log(\Delta p/p)$.

2. Molecular description

The decision rule is to report ''signal-in-the-first-interval'' when

$$\lambda_1+\sum_{j=1}^{n}a_j(L_1+x_{1j}+\Delta L_{1j}+\epsilon_{1j})>\lambda_2+\sum_{j=1}^{n}a_j(L_2+x_{2j}+\Delta L_{2j}+\epsilon_{2j}). \quad (4)$$

The subscripts, 1 and 2, represent the intervals. Peripheral and central internal noise are represented by the respective random variables, ϵ_j and λ, which are assumed to be normally distributed with means of zero and variances that are not necessarily equal across the ϵ_j. The terms serve only as theoretical acknowledgement here. All sources of internal noise will be represented by a single random variable, $\omega=\lambda_2-\lambda_1+\Sigma(\epsilon_{2j}-\epsilon_{1j})$, which is normally distributed with zero mean and a variance of σ_1^2.

3. Decision weights

Decision weights are estimated with a maximum likelihood version of COSS analysis (Berg, 1989). The acronym means ''conditional-on-a-single-stimulus,'' reflecting its origin in experiments involving the sequential presentation of tones. The underlying principle of COSS analysis can be explained by rearranging the terms of Eq. (4), so that

$$x_{1j}-x_{2j}>\frac{\omega+\Sigma_{i\neq j}^{n}a_i(x_{2i}-x_{1i})-\Sigma_{i=1}^{n}a_i(L_2-L_1)}{a_j}. \quad (5)$$

The ΔL terms have been omitted because they have no relevant effect on the analysis. Let the terms on the right side define a normally distributed random variable,[2] Y_j, with a variance of

$$\sigma_{Y_j}^2=\frac{\sigma_\omega^2+2\Sigma_{i\neq j}^{n}a_i^2\sigma_i^2+2(\Sigma_i^n a_i)^2\sigma_L^2}{a_j^2}. \quad (6)$$

The weight of the jth component is a function of $\sigma_{Y_j}^2$. In earlier studies, $\sigma_{Y_j}^2$ was estimated with a least-squares fit of a cumulative Gaussian to a COSS function, which plots the probability of an ''interval-one'' response as a function of the perturbation difference, $x_{1j}-x_{2j}$ (see Berg and Green, 1990). Maximum likelihood estimates of $\sigma_{Y_j}^2$ were obtained here. For each trial-type (i.e., signal in the first or second interval), fixed trials were partitioned into k trials with ''interval-one'' responses and ℓ trials with ''interval-two'' responses. The mean, μ_{Y_j}, and variance, $\sigma_{Y_j}^2$, of the underlying cumulative

Gaussian were found by maximizing the likelihoods,

$$L(\mu_{Y_j},\sigma^2_{Y_j})=\prod_{m=1}^{l}\phi(x_{1jm}-x_{2jm}) \times\prod_{m=1}^{k}[1-\phi(x_{1jm}-x_{2jm})], \tag{7a}$$

and

$$L(\mu_{Y_j},\sigma^2_{Y_j})=\prod_{m=1}^{l}[1-\phi(x_{1jm}-x_{2jm})] \times\prod_{m=1}^{k}\phi(x_{1jm}-x_{2jm}), \tag{7b}$$

with the estimate of $\sigma^2_{Y_j}$ determined by the one yielding the greatest likelihood.

If unit weight is given to the signal component ($j=s$), then the magnitude of weight for the jth component is

$$a_j=\frac{\sqrt{\sigma^2_{Y_s}+2\sigma^2_x}}{\sqrt{\sigma^2_{Y_j}+2\sigma^2_x}}. \tag{8}$$

If Eq. (7a) yields a greater likelihood than Eq. (7b), then a_j is negative; otherwise, the weight is positive. Independent sets of estimates were obtained for each trial type, and the averages comprised the weights used in subsequent calculations. Reliability was assessed by calculating the rms difference between the two sets of independent estimates, a measure referred to as the weight rms.

In theory, $a_L=\Sigma a_j$. However, using the sum of weights as the coefficient for σ^2_I in Eq. (3) introduces a source of inaccuracy because estimation errors are compounded in the sum. Instead, a_L was estimated directly. If Eq. (5) is rewritten so that L_1-L_2 is isolated on the left side, then the sum of terms on the right defines a Gaussian distributed variable, Y_L. Maximum likelihood estimates of the variance, $\sigma^2_{Y_L}$, were obtained by substituting L_1 and L_2 in Eq. (7b). The level weight, a_L, was then obtained by substituting $\sigma^2_{Y_L}+2\sigma^2_L$ in the denominator of Eq. (8).

4. Internal noise

The basis for estimating internal noise can be shown by rearranging Eq. (5),

$$\sum_{j=1}^{n}a_j(x_{1j}-x_{2j})+a_L(L_1-L_2)>\omega, \tag{9}$$

The terms on the left side define the DV. For each of the actual trials, the DV was calculated using the estimated weights of the listener. The results were used to construct a DV function for each trial type, as illustrated in Fig. 1. Least-squares fits of a cumulative Gaussian can be used to estimate σ^2_I. In practice, however, although it is conceptually useful to display a DV function, estimates of σ^2_I were obtained with a maximum likelihood procedure. The procedure was identical to that for estimating decision weights, except that DV was substituted for x_1-x_2 in Eq. (7b). Independent estimates

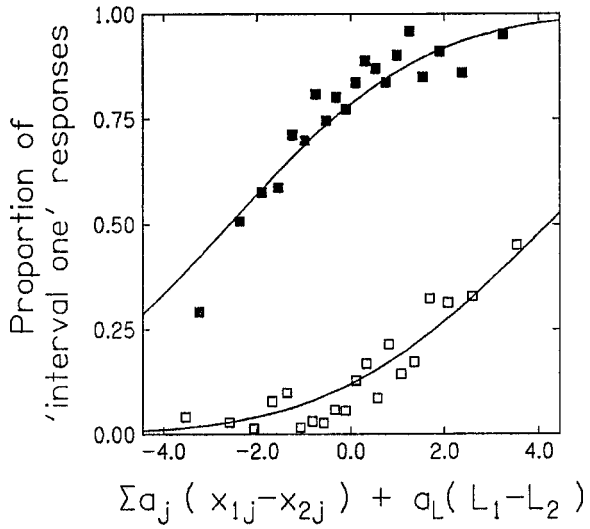


FIG. 1. DV functions. The probability of responding "interval-one" is plotted as a function of the DV and fit with a cumulative Gaussian (from the data of S3). Closed symbols are from trials with the signal in the first interval, with a result of $\sigma_I=2.5$ dB. Open symbols are from trials with the signal in the second interval, with a result of $\sigma_I=1.9$ dB.

of σ^2_I from the two trial-types were averaged to obtain a final estimate.

For the final step in the analysis, estimates of a_j, $j=1,...,n$, a_L, and σ^2_I were used in Eq. (3) to calculate a threshold for the model.

III. RESULTS

Estimates of the decision weights are shown by open symbols in Fig. 2; each is bracketed by two solid symbols representing the independent estimates from each trial type. The level weight, a_L, is represented by the unconnected point on the right. The weight rms is less than 0.1 for four listeners, 0.16 for S5, and 0.42 for S6 (the author). Whereas the expected values of weight estimates are unaffected by internal noise (see Berg, 1989), the variance of estimates is not. The greater weight rms for S6 is consistent with the fact that the highest estimates of internal noise are also found for this listener (see below).

Confidence intervals, derived from computer simulations, were used to infer whether weight estimates are significantly different from zero. For all trials, newly sampled perturbations were substituted for the actual perturbations of the nonsignal components and the data were reanalyzed. This forces a correlation of zero between the perturbations and a listener's responses, and thus decision weights that are truly zero. Confidence intervals around zero are based on the standard deviation of weight estimates, obtained from 100 simulated analyses of the entire data set. Decision weights significantly different from zero are marked by symbols along the bottom of each panel of Fig. 2; solid points denote $p=0.05$ and x's denote $p=0.1$. A test for normality was done using the data of S6; it could not be rejected with a chi-square test ($df=17$; $p=0.05$).

Listener thresholds and model thresholds are listed in Fig. 2. For four listeners, the difference is less than 1 dB. It is of interest to note that S3 and S4 obtain low thresholds by

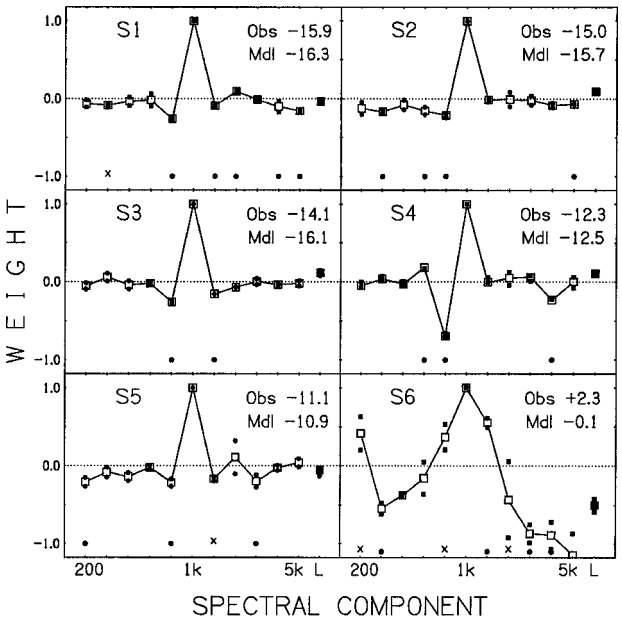


FIG. 2. Decision weights. Connected points in each panel show estimated decision weights; L on the abscissa represents the level weight, a_L. Each point is bracketed by two solid symbols, representing the two independent estimates. Unconnected points along the bottom of each panel mark the weights that are significantly different from zero; solid symbols denote $p=0.05$ and the x's denote $p=0.1$. Observed thresholds and model thresholds are listed in each panel.

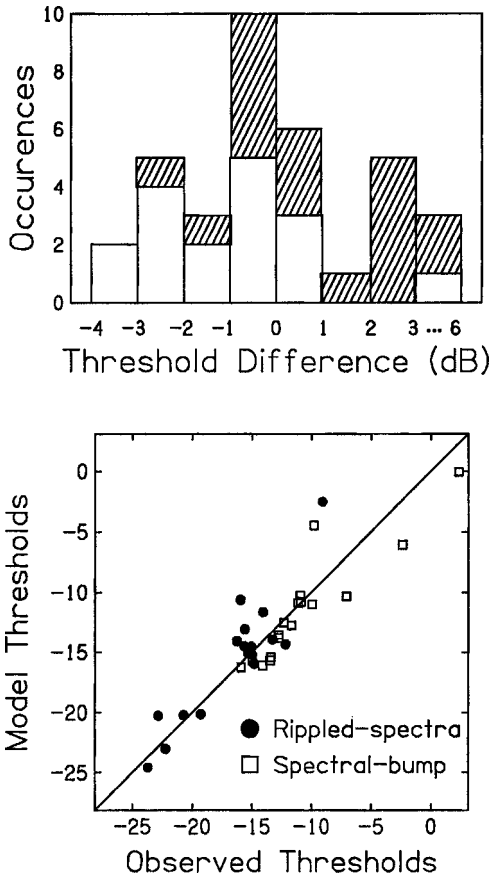


FIG. 3. Histogram and scatterplot of prediction errors. Open bars in the upper panel show prediction errors for the spectral-bump conditions across all three studies and hatched bars show prediction errors for rippled-spectra conditions. The bottom panel shows observed thresholds plotted as a function of model thresholds ($r^2=0.72$).

apparently attending to just a few nonsignal components. On the other hand, as in the case of S6, attending to a greater number of channels does not guarantee good performance.

Estimates of σ_I were in the range of 2 to 3 dB for four listeners, 3.6 dB for S5, and 11.0 dB for S6. Although a formal proof is needed, it is the author's contention that comparisons of σ_I among listeners are permissible. Model thresholds are invariant to multiplicative transformations of the decision weights, and so the constraint of unit weight for the signal component normalizes estimates of σ_I to a common scale with respect to optimal performance.

Data from Berg *et al.* (1992) and Berg and Green (1991) were also analyzed. Berg *et al.* used three-component spectra with a 1-kHz signal component and a bandwidth that varied across conditions; the analysis was completed only for the wideband conditions for which weights for both nonsignal components were negative. Berg and Green used a signal consisting of a sinusoidal ripple across an eight-component spectrum. The results, compiled with current data, are summarized in Fig. 3. The bottom panel shows model thresholds plotted as a function of observed thresholds and the top panel shows the distribution of threshold differences. Data from rippled-spectra and spectral-bump conditions are partitioned in Fig. 3, but there is no significant difference between the distributions. Accuracy of the model is generally within 4.0 dB, a value approximating the standard deviations typically observed for empirical data.

A central issue is the extent to which internal noise is overestimated when optimal weights are assumed. Comparisons between estimates of σ_I obtained by using optimal weights to calculate the DV and those obtained with estimated weights are displayed in Fig. 4. The mean increase in estimates of σ_I when an optimal weight model is used are 6.7 dB for rippled-spectra (upper panel) and 0.4 dB for spectral-bump conditions (lower panel). The optimal weight model yields lower estimates of σ_I in only two cases. For one, the difference is 0.1 dB. For the other, the reliability of weight estimate is questionable because this case also exhibits the greatest weight rms (0.53) across the three studies.

Direct comparisons of internal noise estimates across conditions are inappropriate because σ_I is not invariant to changes in the scale of weights. Multiplying the weights by an arbitrary constant changes σ_I by the same factor, without affecting d'_{obs}. For spectral-bump conditions, as noted above, the constraint of unit weight for the signal component normalizes the scale relative to the ideal observer, thus legitimizing within-condition comparisons. On the other hand, direct comparisons of σ_I for the rippled-spectra conditions cannot be made because the component serving as the anchor is allowed to vary across listeners, and so a maximum weight of unity is not a normalizing constraint as it is for spectral-bump conditions. However, the efficiency measure, η_N, is scale invariant, which allows comparisons both within and between conditions. Mean estimates of η_N are 0.193 for spectral-bump conditions and 0.073 for rippled-spectra conditions. A *posthoc* analysis shows that the difference is significant, $t(34)=5.61$, $p<0.01$. A possible explanation for the

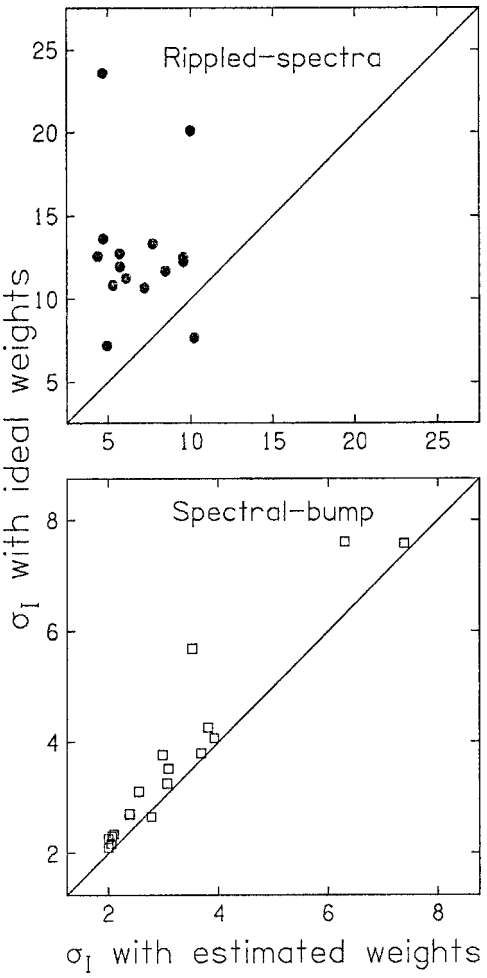


FIG. 4. Comparisons of internal noise estimates obtained with and without optimal decision weights. Estimates of σ_{I} obtained with the DV calculated using optimal weights are plotted as a function of estimates of σ_{I} obtained with the DV calculated using estimated weights.

difference is that it may be more difficult to maintain a consistent weighting strategy in rippled-spectra conditions, an inefficiency that would be manifested as internal noise.

The overall mean prediction error for the model is 0.01 dB. For an optimal weight model the mean error is −0.8 dB. The similarity is more fortuitous than meaningful. Increases in σ_{I} with optimal weights are apparently balanced by zeroing the effects of the level variance (i.e., $a_L=0$). In addition, the summed signal (i.e., $\Sigma a_i\Delta_i$) for rippled-spectra is generally greater with optimal weights.

IV. DISCUSSION

Establishing concordance with standard psychophysical measurement is an important step in validating molecular analysis. This study demonstrates that such an analysis provides an accurate description of profile discrimination thresholds. What is most valuable about molecular analysis, however, is not the power to predict thresholds, but rather the descriptions and insights about listening behavior that are beyond the purview of conventional psychophysical methods. Regarding tone-in-noise detection, for instance, Watson (1962) first discovered violations of a single-channel energy detector with a molecular analysis, a finding later confirmed with other molecular techniques (Pfafflin and Mathews, 1966; Ahumada and Lovell, 1971; Gilkey and Robinson, 1986; Richards and Buss, 1996). In a task involving the discrimination of three-tone, narrow-band spectra, Berg *et al.* (1992) and Southworth and Berg (1995) proposed that different patterns of decision weights correspond to differences in the cue used to make the discriminations (i.e., pitch, loudness, roughness). Consistent agreement was found between inferences drawn from the molecular analysis and the self-reports of listeners. The significance of this finding is underscored by considering that the extreme variation in self-reports would likely be viewed with some incredulity if interpretations of the data were based solely on thresholds.

The current study presents another facet of molecular psychophysics, namely a more accurate means of estimating internal noise. An obvious generalization is multiple observation tasks in which deviations from the ''square-root-of-*n* rule'' are used to derive estimates of internal noise (e.g., Swets *et al.*, 1959). The rate at which d'_{obs} reaches an asymptote as the number of observations increases is thought to be determined by a combination of peripheral and central internal noise (Robinson and Berg, 1986; Lutfi 1989; Weber, 1995). An additional factor that might affect the shape of sensitivity functions is an increasing inability of a listener to maintain a uniform set of weights as the number of observations increases. Systematic inefficiencies, undetected by a molar analysis, would likely manifest as internal noise. Molecular techniques correct such estimates by factoring the listener's decision weights into the analysis.

Molecular psychophysical techniques also offer demanding tests of models. It may often be the case that two models provide equally good predictions of threshold data. One might be superior, however, in its capacity to explain the microstructure of data. Another attribute is that the degrees of freedom are essentially limitless with a molecular analysis, thus allowing tests of increasingly complex models. With respect to model development, DV functions appear to be sensitive enough to fine-tune the DV of a model. This can be demonstrated by comparing the current model to one which disregards information about absolute intensity. With $a_L=0$, estimates of σ_{I} increase by an average of 0.73 dB. Statistical significance of this seemingly small value is assured by the fact that an increase is found for all 36 cases across the three experiments. The finding implies that listener's decisions are affected by absolute intensity, which in turn suggests that the process of level normalization is imperfect.

The fact that estimates of σ_{I} change when the DV is modified compels a reconceptualization of what is meant by ''internal noise.'' There is certainly some inherent biological randomness that denotes the limits of explanation. Just as likely is the existence of undiscovered systematic inefficiencies, that part which is potentially explainable, but is instead manifested as internal noise. Progressing from an optimal weight version of the current model, to a version using estimated weights, but with $a_L=0$, and finally to the full model, we find that each modification accounts for a greater proportion of the variance, as indicated by a decrease in σ_{I}. If sequential response dependencies were factored into the DV,

further decreases in estimates of σ_I might be expected. It may also be the case that nonlinear DVs (Lutfi, 1992b; Dai *et al.*, 1996; Richards, 2002) lead to a further reduction in estimates of σ_I, although the development of nonlinear models will best be approached with a different stimulus configuration because the current fits leave little room for improvement.

ACKNOWLEDGMENTS

I would like to thank Donald Robinson and David Green for many fruitful discussions. Robert Lutfi, R. Duncan Luce, and Geoffry Iverson provided comments on early versions of the paper. Quang Nguyen assisted with data collection. This work was supported by grants from the ONR and AFOSR.

[1]COSS is an acronym for ''conditional-on-a-single-stimulus,'' in reference to a psychometric function which plots the probability of a response (e.g., probability of responding ''signal in first interval'') as a function of the stimulus value of a single component of a multicomponent stimulus.

[2]Deviations from normality occur because the distribution of L is uniform, although they are thought to be slight. The degree of deviation is a function of the level weight, a_L, which is generally close to zero.

Ahumada, A. (**1967**). ''Detection of tones masked by noise: A comparison of human observers with digital-computer energy detectors of varying bandwidths,'' Technical Report No. 29, Human Communication Laboratory, UCLA.

Ahumada, Jr., A., and Lovell, J. (**1971**). ''Stimulus features in signal detection,'' J. Acoust. Soc. Am. **49**, 1751–1756.

Ahumada, Jr., A., Marken, R., and Sandusky, A. (**1975**). ''Time and frequency analysis of auditory signal detection,'' J. Acoust. Soc. Am. **57**, 385–390.

Arnold, S. F. (**1981**). *The Theory of Linear Models and Multivariate Analysis* (Wiley, New York).

Berg, B. G. (**1987**). ''Internal noise in auditory decision tasks,'' unpublished doctoral dissertation, Indiana University, Bloomington, IN.

Berg, B. G. (**1989**). ''Analysis of weights in multiple observation tasks,'' J. Acoust. Soc. Am. **86**, 1743–1746.

Berg, B. G. (**1990**). ''Observer efficiency and weights in a multiple observation task,'' J. Acoust. Soc. Am. **88**, 149–158.

Berg, B. G., and Green, D. M. (**1990**). ''Spectral weights in profile listening,'' J. Acoust. Soc. Am. **88**, 758–766.

Berg, B. G., and Green, D. M. (**1991**). ''Discrimination of complex spectra: spectral weights and performance efficiency,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 373–379.

Berg, B. G., and Robinson, D. E. (**1987**). ''Multiple observations and internal noise,'' J. Acoust. Soc. Am. **81**, S33.

Berg, B. G., Nguyen, Q., and Green, D. M. (**1992**). ''Discrimination of narrow band spectra: I. Spectral weights and pitch cues,'' J. Acoust. Soc. Am. **92**, 1911–1918.

Bernstein, L. R., and Green, D. M. (**1987**). ''Detection of simple and complex changes of spectral shape,'' J. Acoust. Soc. Am. **82**, 1587–1592.

Dai, H. (**1994**). ''Signal-frequency uncertainty in spectral-shape discrimination: Psychometric functions,'' J. Acoust. Soc. Am. **96**, 1388–1396.

Dai, H., and Berg, B. G. (**1992**). ''Spectral-temporal weights in a profile task,'' J. Acoust. Soc. Am. **92**, 1346–1355.

Dai, H., Nguyen, Q., and Green, D. M. (**1996**). ''Decision rules of listeners in spectral-shape discrimination with or without signal frequency uncertainty,'' J. Acoust. Soc. Am. **99**, 2298–2306.

Doherty, K. A., and Lutfi, R. A. (**1996**). ''Spectral weights for overall level discrimination in listeners with sensorineural hearing loss,'' J. Acoust. Soc. Am. **99**, 1053–1058.

Doherty, K. A., and Lutfi, R. A. (**1999**). ''Level discrimination of single tones in a broadband signal by normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **105**, 1831–1840.

Durlach, N. I., Braida, L. D., and Ito, Y. (**1986**). ''Towards a model for discrimination of broadband signals,'' J. Acoust. Soc. Am. **80**, 63–72.

Gilkey, R. H., and Robinson, D. E. (**1986**). ''Models of auditory masking: A molecular psychophysical approach,'' J. Acoust. Soc. Am. **79**, 1499–1510.

Green, D. M. (**1964**). ''Consistency of auditory detection judgments,'' Psychol. Rev. **71**, 392–407.

Green, D. M. (**1988**). *Profile Analysis: Auditory Intensity Discrimination* (Oxford U.P., New York).

Green, D. M., and Berg, B. G. (**1991**). ''Spectral weights and the profile bowl,'' Q. J. Exp. Psychol. A **43A**(3), 449–458.

Green, D. M., and Mason, C. R. (**1985**). ''Auditory profile analysis: Frequency, phase and Weber's law,'' J. Acoust. Soc. Am. **77**, 1155–1161.

Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Lentz, J. J., and Richards, V. M. (**1997**). ''Sensitivity to changes in overall level and spectral shape: An evaluation of a channel model,'' J. Acoust. Soc. Am. **101**, 3625–3635.

Lentz, J. J., Richards, V. M., and Matiasek, M. M. (**1999**). ''Different auditory filter bandwidth estimates based on profile analysis, notched noise, and hybrid tasks,'' J. Acoust. Soc. Am. **106**, 2779–2792.

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Lutfi, R. A. (**1989**). ''Informational processing of complex sound. I: Intensity discrimination,'' J. Acoust. Soc. Am. **86**, 934–944.

Lutfi, R. A. (**1990**). ''Informational processing of complex sound. II: Cross-dimensional analysis,'' J. Acoust. Soc. Am. **87**, 2141–2148.

Lutfi, R. A. (**1992a**). ''Informational processing of complex sounds. III: Interference,'' J. Acoust. Soc. Am. **91**, 3391–3401.

Lutfi, R. A. (**1992b**). ''Comment on 'Analysis of weights in multiple observation tasks' [J. Acoust. Soc. Am. **86**, 1743–1746 (**1989**)],'' J. Acoust. Soc. Am. **91**, 507–508.

Lutfi, R. A. (**1995**). ''Correlation coefficients and correlation ratios as estimates of observer weights in multiple observation tasks,'' J. Acoust. Soc. Am. **97**, 1333–1334.

Mahalanobis, P. C. (**1936**). ''On the generalized distance in statistics,'' Proc. Nat. Inst. Sci. Ind. **12**, 49–55.

Montgomery, D. A. (**1991**). ''Information selection under processing limitations: A weight analysis,'' unpublished Master's thesis, Indiana University, Bloomington, IN.

Murray, R. F., Bennett, P. J., and Sekuler, A. B. (**2002**). ''Optimal methods for calculating classification images: Weighted sums,'' J. Vision **2**, 79–104.

Pfafflin, S. M., and Mathews, M. V. (**1966**). ''Detection of auditory signals in reproducible noise,'' J. Acoust. Soc. Am. **39**, 340–345.

Richards, V. M. (**2002**). ''Effects of a limited class of nonlinearities on estimates of relative weights,'' J. Acoust. Soc. Am. **111**, 1012–1017.

Richards, V. M., and Buss, E. (**1996**). ''Frequency correlation functions for the detection of a tone added to modulated noise maskers,'' J. Acoust. Soc. Am. **99**, 1645–1652.

Richards, V. M., and Zhu, S. (**1994**). ''Relative estimates of combination weights, decision criteria and internal noise based on correlation coefficients,'' J. Acoust. Soc. Am. **95**, 423–434.

Robinson, D. E., and Berg, B. G. (**1986**). ''A partitioned variance model for multiple observations,'' paper presented at the 19th annual meeting of the Society for Mathematical Psychology, Cambridge, MA.

Sadralodabai, T., and Sorkin, R. D. (**1999**). ''Effects of temporal position, proportional variance, and proportional duration on decision weights in temporal pattern discrimination,'' J. Acoust. Soc. Am. **105**, 358–365.

Sorkin, R. D. (**1990**). ''Perception of temporal patterns defined by tonal sequences,'' J. Acoust. Soc. Am. **87**, 1695–1701.

Sorkin, R. D., Robinson, D. E., and Berg, B. G. (**1987**). ''A detection theory method for the analysis of auditory and visual displays,'' in *Proceedings of the 31st Annual Meeting of the Human Factors Society*, pp. 1184–1188.

Sorkin, R. D., Mabry, T. R., Weldon, M. S., and Elvers, G. (**1991**). ''Integration of information from multiple element displays,'' Org. Behav. Human Decision Process. **49**, 167–187.

Southworth, C., and Berg, B. G. (**1995**). ''Multiple cues for the discrimination of narrow-band sounds,'' J. Acoust. Soc. Am. **98**, 2486–2492.

Spiegel, M. F., and Green, D. M. (**1981**). ''Two procedures for estimating internal noise,'' J. Acoust. Soc. Am. **70**, 69–73.

Swets, J. A., Shipley, E. F., McKey, M. J., and Green, D. M. (**1959**). ''Multiple observation of signals in noise,'' J. Acoust. Soc. Am. **31**, 514–521.

Tanner, W. P., and Birdsall, T. G. (**1958**). ''Definitions of d' and ν as psychophysical measures,'' J. Acoust. Soc. Am. **30**, 922–928.

Treisman, M., and Williams, T. (**1984**). ''A theory of criterion setting with an application to sequential dependencies,'' Psychol. Rev. **91**, 68–111.

Watson, C. S. (**1962**). ''Signal detection and certain physical characteristics of the stimulus during the observation interval,'' unpublished doctoral dissertation, Indiana University, Bloomington, IN.
Watson, C. S., Franks, J. R., and Hood, D. C. (**1972**). ''Detection of tones in the absence of external masking noise: I. Effects of signal intensity and signal frequency,'' J. Acoust. Soc. Am. **52**, 633–643.
Weber, D. L. (**1995**). ''Information integration of pooled versus paired samples,'' J. Acoust. Soc. Am. **98**, 1903–1908.

Hands help hearing: Facilitatory audiotactile interaction at low sound-intensity levels

Martin Schürmann,[a] Gina Caetano, Veikko Jousmäki, and Riitta Hari
Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, FIN-02015 HUT, Espoo, Finland

(Received 8 May 2003; revised 25 September 2003; accepted 6 November 2003)

Auditory and vibrotactile stimuli share similar temporal patterns. A psychophysical experiment was performed to test whether this similarity would lead into an intermodal bias in perception of sound intensity. Nine normal-hearing subjects performed a loudness-matching task of faint tones, adjusting the probe tone to sound equally loud as a reference tone. The task was performed both when the subjects were touching and when they were not touching a tube that vibrated simultaneously with the probe tone. The subjects chose on average 12% lower intensities ($p<0.01$) for the probe tone when they touched the tube, suggesting facilitatory interaction between auditory and tactile senses in normal-hearing subjects. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639909]

PACS numbers: 43.66.Lj, 43.66.WV [MRL] Pages: 830–832

Persons with hearing impairment may perceive sounds (including speech) using their sense of touch, either relying on touch alone or on touch combined with aided audition (Gault, 1926; Sherrick, 1984; Weisenberger and Miller, 1987; Lynch *et al.*, 1988; Levänen and Hamdorf, 2001). These findings suggest some shared neural substrates for auditory and tactile perception. Accordingly, a ''crosstalk'' between auditory and tactile modalities can be demonstrated even in normal-hearing subjects, although such an interaction in everyday life mostly goes unnoticed. For example, tactile exploration of surfaces elicits auditory and tactile input, but the percept is typically dominated by the tactile component (Lederman, 1979). Likewise, when subjects rub their palms together, the tactile input dominates over the concomitant sound. However, as soon as the high frequencies of that sound are accentuated, subjects report a modified tactile percept (the ''parchment-skin illusion,'' Jousmäki and Hari, 1998).

Here, we quantified audiotactile interaction in normal-hearing subjects whom we asked to adjust the intensity of probe tones to sound equally loud as a low-intensity reference tone. When the subjects touched a tube that vibrated in synchrony with the probe tones, they chose lower tone intensities than without the vibration, suggesting facilitatory audiotactile interaction.

We tested nine subjects (five females, four males; 24–41 years, median 27 years; all but one right-handed, and all normal-hearing by self-report) after informed consent. Pairs of 200-Hz tones (a constant-intensity reference tone of 900 ms and an adjustable probe tone of 500 ms, 100-ms pause in-between; see Fig. 1) were presented binaurally via headphones once every 2 s. The subject was instructed to adjust the probe and reference tones to sound equally loud.

The tones were embedded within a continuous masker (white noise ~60 dB above hearing threshold on one of the authors: constant for all subjects). For each subject, the intensity of the reference tone was adjusted to a level of 10 dB above the individual threshold for detecting the tone within the masking white noise.

The reference–probe pairs were presented in combination with fixed-intensity 200-Hz vibrations, delivered via a vibrating tube simultaneously with the adjustable probe tone. In the ''sound & touch'' condition, the subject's left-hand fingers were in contact with the vibrating tube. In the ''sound only'' condition, the subject did not touch the vibrating tube. Touching the tube resulted in a weak percept of vibration in the fingers. The vibration was equally strong for all subjects: 24–28 dB above tactile threshold, as tested in six subjects separately from the main experiment, with dB values calculated from voltages at tube input.

All subjects were trained to touch the tube in a similar manner (at a marked area, with fingers rather than palm, and without squeezing the tube). In an additional measurement on one of the authors, the perceived vibration varied maximally by 4 dB from gentle to firm touch, a range of grip force much wider than applied in the main experiment. The noise played via the headphones was effective in masking any tube-produced sounds, as was established in pilot experiments on two of the authors.

The subjects first practiced the loudness matching task for 10–20 min and then adjusted the probe during ten sound only and ten sound & touch conditions, presented in an alternating order (sound only, sound & touch, sound only,..., sound & touch) to minimize adaptation of Pacinian corpuscles to the tactile stimuli. The experimenter measured the adjusted probe amplitude (in rms) as soon as the subject signaled completion of the task. After each trial, the experimenter changed the probe tone to an arbitrary loudness level, either above or below the reference tone. Pauses were given when needed. Depending on the subject, the experiment lasted for 30–90 min.

In the whole group of nine subjects, the adjusted probe tone intensities were weaker in the sound & touch than in the sound only condition (mean difference of the individual median values $-12\pm4\%$, $p=0.007$, Wilcoxon's signed rank

[a]Author to whom correspondence should be addressed. Electronic mail: martins@neuro.hut.fi

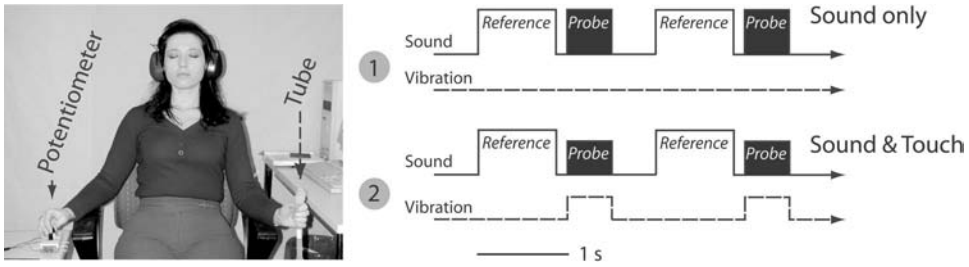


FIG. 1. Left panel: Subject is listening to tones delivered via headphones and simultaneously controlling sound-intensity level with a multiturn potentiometer. Vibratory stimuli are delivered to the subject's left-hand fingers via a blind-ended silicon tube (diameter 16 mm, wall thickness 2 mm) attached to a purpose-built stimulator (loudspeaker in a funnel-shaped housing, not shown, resembling that used by Levänen *et al.*, 1998). Right panel: Pairs of 200-Hz tones (constant-intensity Reference and adjustable Probe), occurring once every 2 s, were embedded within masking white noise and presented via headphones. The 200-Hz vibratory stimuli were driven by the same electrical signals as the probe tones but they were always of constant intensity. The subject touched the tube in the ''Sound & Touch'' condition but did not touch it in the ''Sound only'' condition; in both situations the tube, however, vibrated in the same way.

test for paired samples). This difference between the conditions was statistically significant also at individual level in seven out of nine subjects ($p<0.05$, Wilcoxon's rank sum test for independent samples; see Fig. 2). Visual inspection of the data suggested in the adjusted probe tone intensities a weak but statistically nonsignificant tendency to decrease from the first to the last trial (mean decrease 4% as estimated from linear regression; equal trend for sound only and sound & touch conditions).

These data suggest facilitatory audiotactile interaction in normal-hearing subjects when they listen to low-intensity tones embedded within noise. The results agree with tactile input to the auditory cortex, demonstrated with magnetoencephalographic recordings in a deaf human subject (Levänen *et al.*, 1998) and with intracranial recordings in monkeys (Schroeder *et al.*, 2001). More extensive brain-imaging studies are, however, needed to identify the sites of audiotactile interaction in the human brain (Calvert, 2000; Foxe *et al.*, 2002; Lütkenhöner *et al.*, 2002; Gobbelé *et al.*, 2003), and in deciding whether the interaction between sensory modalities takes place at the perceptual or decision level (Massaro, 1999; Vroomen and de Gelder, 2000)

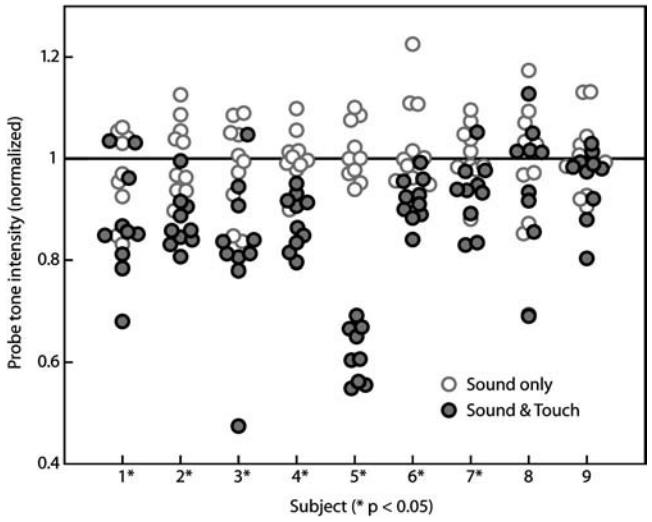


FIG. 2. Probe-rms amplitudes for all nine subjects in all ten ''sound only'' and ten ''sound & touch'' trials; the amplitudes were normalized so that each subject's median amplitude in ''sound only'' trials equals 1.0. The seven subjects in whom the differences between the conditions were statistically significant are marked with asterisks.

In everyday life, audiotactile interaction is rarely noticed, but some illusions resulting from modification of the relative saliences of the stimuli may make the interaction evident (e.g., ''parchment-skin illusion;'' Jousmäki and Hari, 1998; Guest *et al.*, 2002). Additional evidence of audiotactile interaction derives from ''tactile capture of audition,'' in which lateralized sounds can be mislocalized when concomitant tactile stimuli are presented to body midline (Caclin *et al.*, 2002). Moreover, subjects with tactile deficits, e.g., patients using a hand prosthesis after hand amputation, benefit from auditory feedback during tactile exploration (Lundborg and Rosen, 2001). Similarly, in a virtual-reality setup subjects may learn faster to identify texture by touch or may perceive a better quality of match to a test material when complementary auditory information or cues are supplied (Hendrix *et al.*, 1999; Lederman *et al.*, 2003). Thus, auditory input may be important for tactile perception even when it does not reach awareness. In the present study, the low intensities of both auditory and tactile stimuli probably improved our possibilities to detect and quantify the audiotactile interaction in normal-hearing subjects.

ACKNOWLEDGMENTS

We are grateful to Ronny Schreiber for technical assistance. Supported by Academy of Finland, Sigrid Jusélius Foundation, European Union Large-Scale Facility Neuro-BIRCH III at the Brain Research Unit of the Low Temperature Laboratory, Helsinki University of Technology, and Fundação para a Ciência e Tecnologia Programa Praxis XXI, Ministério da Ciência e do Ensino Superior, Portugal.

Caclin, A., Soto-Faraco, S., Kingstone, A., and Spence, C. (**2002**). ''Tactile 'capture' of audition,'' Percept. Psychophys. **64**, 616–630.
Calvert, G. (**2000**). ''Crossmodal processing in the human brain: Insights from functional neuroimaging studies,'' Cereb. Cortex **11**, 1110–1123.
Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., Ritter, W., and Murray, M. M. (**2002**). ''Auditory–somatosensory multisensory processing in auditory association cortex: an fMRI study,'' J. Neurophysiol. **88**, 540–543.
Gault, R. H. (**1926**). ''Touch as a substitute for hearing in the interpretation and control of speech,'' Arch. Otolaryngol. **3**, 121–135.

Gobbelé, R., Schürmann, M., Forss, N., Juottonen, K., Buchner, H., and Hari, R. (**2003**). ''Activation of the human parietal and temporoparietal cortices during audiotactile interaction,'' Neuroimage **50**, 503–511.
Guest, C., Catmur, C., Lloyd, D., and Spence, C. (**2002**). ''Audiotactile interactions in roughness perception,'' Exp. Brain Res. **146**, 161–171.
Hendrix, C., Cheng, P., and Durfee, W. (**1999**). ''Relative influence of sensory cues in a multimodal environment,'' Proceedings of the ASME Dynamics Systems and Control Division **63**, 59–64.
Jousmäki, V., and Hari, R. (**1998**). ''Parchment-skin illusion: Sound-biased touch,'' Curr. Biol. **8**, R190.
Lederman, S. J. (**1979**). ''Auditory texture perception,'' Perception **8**, 93–103.
Lederman, S. J., Martin, A., Tong, C., and Klatzky, R. L. (**2003**). ''Relative performance using haptic and/or touch-produced auditory cues in a remote absolute texture identification task,'' 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS '03), Los Angeles, CA.
Levänen, S., and Hamdorf, D. (**2001**). ''Feeling vibrations: Enhanced tactile sensitivity in congenitally deaf humans,'' Neurosci. Lett. **301**, 75–77.
Levänen, S., Jousmäki, V., and Hari, R. (**1998**). ''Vibration-induced auditory cortex activation in a congenitally deaf adult,'' Curr. Biol. **8**, 869–872.
Lundborg, G., and Rosen, B. (**2001**). ''Sensory substitution in prosthetics,'' Hand Clin. **17**, 481–488.
Lütkenhöner, B., Lammertmann, C., Simões, C., and Hari, R. (**2002**). ''Magnetoencephalographic correlates of audiotactile interaction,'' Neuroimage **15**, 509–522.
Lynch, M. P., Eilers, R. E., Oller, D. K., and Lavoie, L. (**1988**). ''Speech perception by congenitally deaf subjects using an electrocutaneous vocoder,'' J. Rehabil. Res. Dev. **25**, 41–50.
Massaro, D. W. (**1999**). ''Speechreading: Illusion or window into pattern recognition,'' Trends Cogn. Sci. **3**, 310–317.
Schroeder, C. E., Lindsley, R. W., Specht, C., Marcovici, A., Smiley, D. F., and Javitt, D. C. (**2001**). ''Somatosensory input to auditory association cortex in the macaque monkey,'' J. Neurophysiol. **85**, 1323–1327.
Sherrick, C. E. (**1984**). ''Basic and applied research on tactile aids for deaf people: Progress and prospects,'' J. Acoust. Soc. Am. **75**, 1325–1342.
Vroomen, J., and de Gelder, B. (**2000**). ''Crossmodal integration: A good fit is no criterion,'' Trends Cogn. Sci. **4**, 37–38.
Weisenberger, J. M., and Miller, J. D. (**1987**). ''The role of tactile aids in providing information about acoustic stimuli,'' J. Acoust. Soc. Am. **82**, 906–916.

The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer[a)]

Monica L. Hawley[b] and Ruth Y. Litovsky[c]
Hearing Research Center and Department of Biomedical Engineering, Boston University, Boston, MA 02215

John F. Culling
School of Psychology, Cardiff University, P.O. Box 901, Cardiff, CF10 3YG, U.K.

(Received 18 January 2003; revised 14 November 2003; accepted 17 November 2003)

The ''cocktail party problem'' was studied using virtual stimuli whose spatial locations were generated using anechoic head-related impulse responses from the AUDIS database [Blauert *et al.*, J. Acoust. Soc. Am. **103**, 3082 (1998)]. Speech reception thresholds (SRTs) were measured for Harvard IEEE sentences presented from the front in the presence of one, two, or three interfering sources. Four types of interferer were used: (1) other sentences spoken by the same talker, (2) time-reversed sentences of the same talker, (3) speech-spectrum shaped noise, and (4) speech-spectrum shaped noise, modulated by the temporal envelope of the sentences. Each interferer was matched to the spectrum of the target talker. Interferers were placed in several spatial configurations, either coincident with or separated from the target. Binaural advantage was derived by subtracting SRTs from listening with the ''better monaural ear'' from those for binaural listening. For a single interferer, there was a binaural advantage of 2–4 dB for all interferer types. For two or three interferers, the advantage was 2–4 dB for noise and speech-modulated noise, and 6–7 dB for speech and time-reversed speech. These data suggest that the benefit of binaural hearing for speech intelligibility is especially pronounced when there are multiple voiced interferers at different locations from the target, regardless of spatial configuration; measurements with fewer or with other types of interferers can underestimate this benefit.
[DOI: 10.1121/1.1639908]

PACS numbers: 43.66.Pn, 43.71.Gv [PFA] Pages: 833–843

I. INTRODUCTION

In many social situations, listeners receive simultaneous sounds from many sources. Perceptually segregating a single target voice from a competing milieu, so that it can be individually understood, has been termed ''the cocktail-party problem'' (Cherry, 1953). A number of cues and processes that contribute to the solution of the cocktail-party problem have been identified. There are four that are of particular relevance to the current study.

First, spatially separating the target and interferers improves understanding of the target speech. In the free field or ''virtual free-field,'' the effect is known as ''spatial release from masking'' (Plomp and Mimpen, 1981; Bronkhorst and Plomp, 1992; Nilsson *et al.*, 1994; Koehnke and Besing, 1996; Peissig and Kollmeier, 1997; Hawley *et al.*, 1999; Shinn-Cunningham *et al.*, 2001; Litovsky *et al.*, 2002). Spatial release from masking can be regarded as having two components (Durlach, 1963; vom Hövel, 1984; Zurek, 1992; Bronkhorst, 2000): monaural advantage arises directly from improvements in signal-to-noise ratio at the ''best'' ear (BE), which are caused by headshadow; binaural advantage arises from binaural unmasking (BU) of the low-frequency parts of the speech signal, which are largely facilitated by differences in interaural time delay (ITD) between competing sources (Zurek, 1992; Bronkhorst and Plomp, 1988; Durlach, 1963; Culling and Summerfield, 1995; Breebaart *et al.*, 2001a,b,c). This BE+BU account is distinct from that provided by auditory scene analysis (Bregman, 1990), which suggests that spatial release from masking involves the grouping of sound elements from one direction and segregation of that group from elements of interfering sound in different directions. The BE+BU interpretation separates the roles of ITDs and headshadow, while, in the auditory scene analysis, both contribute to the initial determination of sound direction. The present study attempts to differentiate between these accounts by comparing monaural and binaural performance in a variety of listening situations.

Second, understanding of the target speech depends upon the temporal properties of the interfering sound. A speech interferer has a fluctuating frequency spectrum and amplitude envelope. In contrast, speech-shaped noise has a long-term spectrum which matches that of speech, but lacks such modulation (e.g., MacKeith and Coles, 1971; Plomp and Mimpen, 1979; Festen, 1993; Koehnke and Besing, 1996). The effect of the temporal envelope can be investigated using speech-modulated noise, whose temporal envelope is also derived from speech. Dips in the temporal envelope of the interferer are beneficial to understanding of the target voice, presumably due to the transitory improvement

[a)]Portions of this paper were presented at the 137th Meeting of the Acoustical Society of America, March, 1999, British Society of Audiology 2000, and the Meeting of the Association for Research in Otolaryngology, February, 2000.
[b]Current address: Dept. of Otolaryngology, University of Maryland Medical School, 16 S. Eutaw St., Suite 500, Baltimore, MD 21201.
[c]Current address: University of Wisconsin Waisman Center, 1500 Highland Ave., Madison WI 53705.

of signal-to-noise ratio (Festen and Plomp, 1990).

Third, differences in fundamental frequency (F0) between concurrent voices enable listeners to better understand those voices (Brokx and Nooteboom, 1982). Experiments with simultaneous vowels have shown this improvement in understanding to be dependent upon the harmonic structure of the interfering sound, rather than that of the target sound. Lea (1992) found that if one vowel in a pair was noise-excited, detection of the noise-excited and not the harmonic vowel improved. Similarly, Summerfield and Culling (1992) and de Cheveigné *et al.* (1995) found that if one vowel in a pair was inharmonic, identification of this vowel improved (compared to the same-F0 or both-inharmonic cases) and not that of the harmonic one. These data are consistent with the idea that the interfering source is perceptually cancelled (de Cheveigné, 1997). Therefore, when a speech interferer is replaced by speech-modulated noise there can be no advantage from F0 differences. A similar effect may be expected when multiple interferers are presented, although this would depend upon whether the putative canceling mechanism can recursively cancel multiple F0's. We are not aware of any direct perceptual evidence on the effect of multiple F0's among the interferers. These experiments have usually involved stimuli with static fundamental frequencies, but some studies (e.g., Darwin and Culling, 1990; Summerfield, 1992; Culling *et al.*, 1994) have employed modulated F0's, and their results suggest that listeners can exploit instantaneous differences in F0 as proficiently as sustained ones. Thus, when an interfering voice has the same mean F0 (as, for instance, when it is a recording of the same individual), natural modulation of the voice will introduce instantaneous differences in F0 that listeners can exploit. As an illustration of this point, we used Praat to measure the F0 in semitones of each of the voices used in the present study for all of the available recordings of their voices and for every analysis frame. We then calculated the variance of each. The mean instantaneous difference in F0 between randomly selected frames of the same voice can be predicted from the variance sum law; it is $\sqrt{2\sigma_{F0}^2}$, where σ_{F0}^2 is the variance of the voice F0. The values we derived in this way were 5.6 semitones for one voice (known as "DA") and 4.5 semitones for the other ("CW").

Fourth, the interfering speech carries linguistic content, which can be confused with the content of the target voice. This confusion can be regarded as a form of "informational masking." Such masking is a disruption of performance that cannot be accounted for by a simple model of energetic masking (i.e., overlap in the frequencies of the target and interferer). Rather, the masker carries some other information regarding the stimuli and listening conditions, which interferes with perception of the target content (Pollack and Pickett, 1958; Lutfi, 1990; Kidd *et al.*, 1998). Most of what is known about informational masking has been investigated using nonspeech stimuli; however, recent studies using speech as both target and interferer suggest that informational masking might play an important role in the cocktail-party problem (Brungart *et al.*, 2001). When a real-speech interferer is replaced by speech-modulated noise, one may expect some advantage to accrue from the removal of this interference. Thus, this effect tends to oppose the effect of losing F0 differences. In order to differentiate these two effects, one can employ a time-reversed speech interferer, which possesses an F0, but lacks recognizable linguistic content above the phoneme level. A time-reversed speech interferer may show some release of informational masking due to the removal of these components of the interferer information.

In summary, there are four main effects that have been studied with respect to the "cocktail party effect," but their relative importance, especially in multi-talker environments, is poorly understood. While many studies have investigated these four effects individually, few have addressed interactions between them. In addition, few paradigms have been extended towards more complex, ecologically relevant situations in which multiple competing sources occur from various directions. The purpose of the present study was to explore the interaction between the number of interfering sounds, the role of BE and BU when the spatial distribution of interfering sounds are manipulated, and the role of spectral, temporal, and linguistic content. The study thus addresses the problem of understanding the more complex listening situations that are routinely encountered in real life.

The most comprehensive study conducted to date on the effects of multiple sources is that of Peissig and Kollmeier (1997). Peissig and Kollmeier used a virtual sound field presentation of a target source directly ahead and one, two, or three interfering sources, consisting of either speech or speech-shaped noise. In each case, they measured speech reception thresholds (SRTs) using a subjective method with one of the interfering sources in each of 17 different directions. Other interfering sources were in fixed positions. They found that (a) speech produced less interference than noise, and (b) spatial release from masking was smaller with speech than with noise for a single interfering source, but was more robust as additional interfering sources were introduced, such that it showed greater spatial release from masking than noise for three interferers. The results raised some interesting questions.

First, a potential problem with the BE+BU view of spatial unmasking is that models of binaural unmasking appear capable of suppressing only a single interfering source direction, whereas cocktail parties are usually populated by multiple, spatially separated, interfering voices (Peissig and Kollmeier, 1997). The reduction in spatial unmasking that occurred when a single noise interferer was replaced by several suggests support for the BE+BU view. On the other hand, the robustness of spatial unmasking for multiple speech interferers suggests that speech may be an exception to this rule. Peissig and Kollmeier (p. 1668) explain the robustness of spatial unmasking for speech interferers in terms of BE+BU by suggesting that modulation in the interfering sources allows the binaural system to switch between different interferers, cancelling whichever is most energetic at a given point in time. This explanation can account for the robustness of performance with multiple speech interferers, which display independent modulations in their temporal envelopes, compared to performance with multiple continuous-noise interferers, which have no modulation. However,

modulation is one of many physical differences between speech and continuous noise. In order to test Peissig and Kollmeiers explanation, therefore, the present study also used multiple speech-modulated noises. These interferers are identical to the speech-shaped noise interferers except for the critical factor of modulation, which is based upon that of the speech interferers. If independent modulation of interfering sources is the critical factor in producing robust spatial unmasking for multiple interferers, these speech-modulated maskers should produce similarly robust unmasking. In addition, while Peissig and Kollmeier's three-interferer spatial configurations always had at least two interferers in different locations, the present study directly contrasts situations in which three interfering sources are spatially separated with situations in which they are spatially coincident.

Second, the exact role of best-ear listening is unclear in Peissig and Kollmeier's study. Ambiguity occurs for two reasons. One is that they did not contrast best-ear performance with binaural performance. The other is that the condition with three interferers always had fixed-position interferers on both the right and left. In the present study, best-ear performance was measured for all conditions and subtracted from binaural performance to yield a measure of binaural advantage. In addition, conditions were included that contrast three interfering sources in the same hemifield, with a condition in which the interferers are distributed in both hemifields.

A final point of difference between our approach and that of Peissig and Kollmeier is that, in their study, speech intelligibility was measured using a subjective method, whereby subjects adjusted the level of the test sentence to that which corresponded to a subjective judgment of 50% intelligibility. This method was justified on the basis that it enabled data to be collected more rapidly and that a close correlation had been observed in previous studies between objective and subjective SRTs. We preferred to measure speech intelligibility under various interfering conditions using a performance measure.

II. METHODS

A. Listeners

A total of 32 paid participants, 18–36 years old, were recruited from the Boston University community (9 males and 23 females); all were native speakers of English with audiometric thresholds at or below 15 dB HL between 250 and 8000 Hz. None of the listeners were familiar with the sentences used in this study.

B. Conditions

Each listener completed testing in three to six sessions of 1.5 to 2 h each. During these sessions they contributed a single SRT in each of 48 conditions (3 numbers of interferers$\times$4 spatial configurations$\times$4 interferer types). Sixteen listeners provided these SRTs with binaural presentation and 16 with monaural presentation, so the monaural and binaural data sets were collected in exactly the same way but from different sets of listeners. Each listener from the monaural condition could be paired with one from the binaural condition, who completed the different conditions of spatial configuration and interferer type in the same order using the same materials.

Different sets of 16 target-sentence lists were used for data collection using different numbers of interferers. To decrease the effect on the thresholds from using different target-sentence lists and to minimize any order effect, a Latin square design was utilized in which each list was paired with each condition only once and each list occupied a particular place in the order only once. Thus, each listener performed one SRT measurement for each condition and using each list. Each number of interferers had a separate Latin square order using a different set of lists.

C. Simulated anechoic space

Anechoic head-related impulse responses (HRIRs) from the HMSIII acoustic manikin and distributed in the AUDIS collection (Blauert *et al.*, 1998) were used to simulate the spatial locations. The stimulus intended for each position was convolved with the set of HRIRs for the left and right ear. All stimuli for each ear were digitally added and presented to the listener through Sennheiser HD433 headphones while they were seated in a double-walled IAC sound-attenuated booth. For the monaural conditions, only the left headphone was stimulated since this was usually the ''better monaural ear'' defined as the ear with the better signal-to-noise ratio; in the majority of simulated configurations the interfering virtual sound sources were situated to the listeners' right, and were therefore less intense at the left than the right ear.

D. Sound sources

The speech tokens were from the Harvard IEEE corpus (Rothauser, 1969). The recordings[1] used were from two male speakers, each contributing half of the sentences. Six of the longest sentences for each talker were reserved for use as interferers to ensure that all targets were shorter than the interferers. The remaining sentences were made into 64 lists of ten sentences each maintaining a single talker for each list. The interferers paired with the target list were from the same talker.

An interferer of each type (speech; reversed speech; speech-shaped noise; speech-modulated, speech-shaped noise) was made based on each of the six interferer sentences. The noise interferers were filtered to match the long-term spectrum of the speech interferers, calculated for each talker separately. The noise samples were cut to the same length as the matching speech interferer and scaled to the same root-mean-square value. For the speech-modulated, speech-shaped noise, the envelope was extracted from the speech interferer and was used to modulate the noise tokens, giving the same coarse temporal structure as the speech. The envelope of running speech was extracted using a method similar to that described by Festen and Plomp (1990), in which a rectified version of the waveform is low-pass filtered. A first-order Butterworth low-pass filter was used with a 3-dB cutoff at 40 Hz. The time-reversed interferer was

TABLE I. Location of interferers.

Interferer location	Front	Left or distributed on both sides	Distributed on right	Together on right
One interferer	0°	−30°	+60°	+90°
Two interferers	0°, 0°	−30°, +90°	+60°, +90°	+90°, +90°
Three interferers	0°, 0°, 0°	−30°, +60°, +90°	+30°, +60°, +90°	+90°, +90°, +90°

speech reversed in time, end to end. Reversed-speech interferers had the same coarse and fine temporal-spectral structure as speech, but no intelligibility.

E. Sound-source locations

The target location was always at the front (0°). There were conditions with one, two, or three interferers, which were all of the same type in a given condition. Up to three interferers were placed either in the front (0°,0°,0°) distributed on both sides (−30°,60°,90°), distributed on the right side (30°,60°,90°), or from the same location on the right side (90°,90°,90°). See Table I for the full specification of these conditions. The level of each interferer was fixed and so the overall level of the interferers was increased as more interferers were added.

F. SRTs

SRTs were measured using a method similar to that developed by Plomp (1986). Listeners were seated in the sound-attenuated booth in front of a terminal screen. A practice SRT with three interferers for each of the interferer types was given at the start of each session to familiarize the subject with the interferer types and the task.

At the start of each SRT measurement, the level of the target was initially very low. The listener heard the same target sentence and interferer combination repeatedly. Each time the listener pressed the return key the same target sentence and interferer combination was replayed, but with the signal-to-interferer ratio increased by 4 dB. When the listener judged they could hear ''more than half'' of the sentence, they typed in their first transcript. From that point on, an SRT was measured using a one-down/one-up adaptive SRT technique targeting 50% correct speech reception (Levitt, 1971).

Correct speech reception was self-assessed by the listener. After listening to each sentence, the listener typed in their transcript. On pressing the return key, the correct target text was also printed on the screen. Each IEEE sentence had five designated key words and these words were in capital letters in the transcript (e.g., The BIRCH CANOE SLID on the SMOOTH PLANKS.). The listener compared the two transcripts and typed in how many key words were correct. The level of the each trial was raised by 2 dB if two or fewer key words were correct and the level was lowered by 2 dB if three or more key words were correct. The entire transaction was logged in a data file and displayed on the experimenter's computer monitor for verification of scoring reliability. The SRT was determined by averaging the level presented on the last eight trials.[2]

In the speech condition, listeners needed to know the text of the interfering sentences because the interferers were from the same voice as the target sentences and in some conditions all sentences were presented from the front location. The texts of any speech interferers were therefore printed on the screen prior to the start of an SRT measurement. The content, number, and locations of the interferers were fixed throughout the run. In conditions that contained a nonspeech interferer, ''unintelligible'' was printed on the screen.

III. RESULTS AND DISCUSSION

The data were analyzed using the assumption that the observed differences between SRTs for different spatial configurations are the result of two independent processes (best-ear listening and binaural advantage) which are additive in decibels. Using these assumptions, the raw SRTs for monaural and binaural listening were used to calculate three additional statistics.

First, the total advantage of separation for each listener in each condition is determined by subtracting the SRT from a given separated condition from that for the corresponding unseparated condition. The advantage of separation for the binaural condition is called the ''total advantage of separation,'' since it contains advantages due to both head shadow (monaural factor) and binaural processing.

Second, the monaural advantage of separation for each listener in each condition (i.e., best-ear listening) is defined as the difference in SRT between each monaural spatially-separated condition and the corresponding unseparated condition.

Third, the binaural advantage is defined as the part of the total advantage that is not accounted for by the monaural advantage. It is obtained by subtracting (in decibels) the monaural advantage from the total advantage of separation. For this purpose the listeners from the monaural and binaural conditions were paired.[3] This difference measure reflects the binaural processing that occurs in different situations, since it is only present when two ears are available and reflects the benefit over listening with just the better monaural ear.

All five measures are discussed below, but statistical analysis is reserved for the derived monaural and binaural advantages of separation. This statistical choice avoided reanalyzing the same data in different ways. The decision to analyze the component advantage of separation is supported by Figs. 1–3, which show that the component effects produce a clearer, more easily interpreted, pattern than the raw data. Scheffé *post hoc* contrasts between means were performed on all significant results from each ANOVA, using

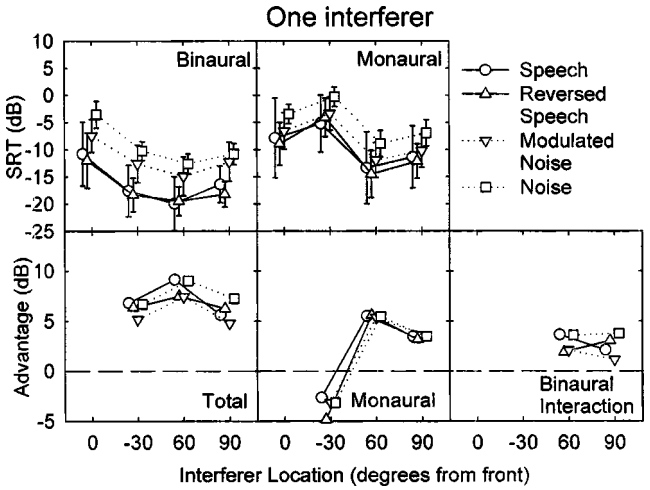


FIG. 1. SRTs and advantages of spatial separation for a single interfering source. The top two panels show means of the raw SRTs, with standard error bars, using two ears (binaural) and using only the left ear (monaural). The lower three panels show the advantage of spatial separation derived by subtracting away the SRT for the nonseparated condition using two ears (total) and using only the left ear (monaural). The binaural advantage is the difference between the total and the monaural advantage.

α=0.05. *Post hoc* one-sample *t*-tests were used to demonstrate deviation of spatial advantages from zero. Bonferroni correction was not used for these *t*-tests because they were intended to identify which spatial advantages differed from zero rather than whether any of them differed.

A. One interferer

The results for a single interferer are shown in Fig. 1.

1. Raw SRTs

For the binaural condition, the SRTs decrease as the interferer location is separated from 0°, the location of the target, regardless of interferer type. The effect of interferer type is seen as an overall shift in the SRTs. The lower SRTs for speech and reversed speech probably reflect the exploitation of differences in F0 between target and interferer (Brokx and Nooteboom, 1982), which may have enabled the interferer to be cancelled (de Cheveigné, 1997).

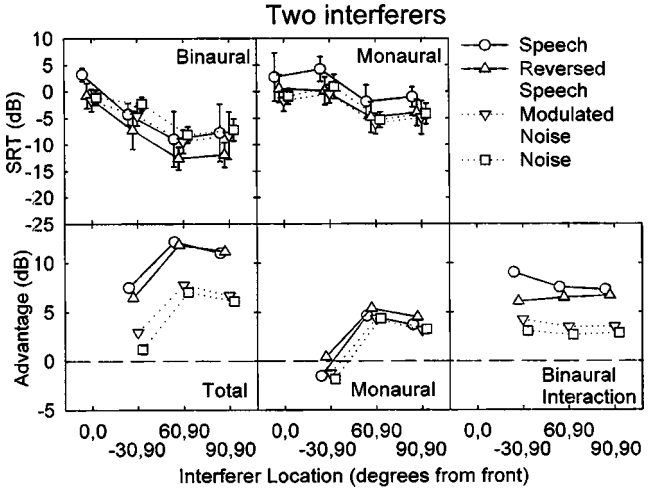


FIG. 2. As in Fig. 1, but for two interfering sources.

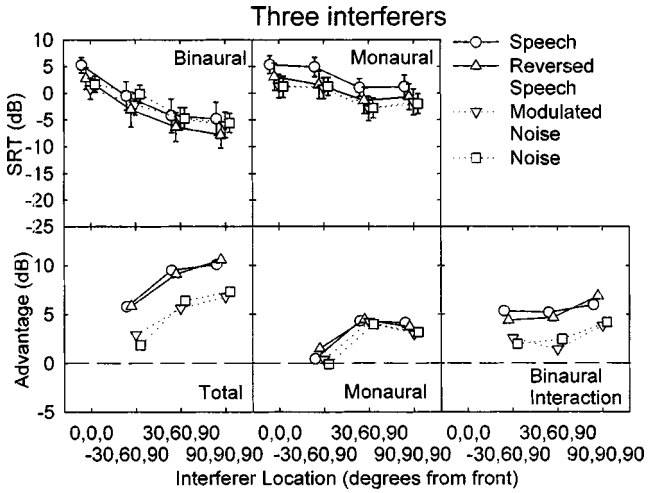


FIG. 3. As in Figs. 1 and 2, but for three interfering sources.

For the monaural condition, the SRTs increase for an interferer at −30° and then fall for interferers at 60° and 90°. The increase at −30° is expected, since, for this interferer location, the left ear is on the same side as the interferer, and so the SNR is not favorable. The ordering of the interferer types is the same as was seen for the binaural condition. However, the difference between the modulated noise interferer and the speech (1.4 dB) and reversed speech (1.9 dB) interferers is not as marked as it was for the binaural condition

2. Advantages of separation

There is a large total advantage (about 6 dB) when moving the interferer 30°, 60°, or 90° from the target location. A similar effect of location is observed using each interferer.

The monaural advantage of separation was negative for an interferer at −30° due to the unfavorable SNR, but +6 and +3 dB for interferers at 60° and 90°, respectively. A two-factor ANOVA (4 interferer types×3 interferer locations) revealed a significant effect of interferer location [$F(2,30)=80$, $p<0.0001$], but not interferer type and no interaction. *Post hoc* analysis of interferer location revealed that all levels of interferer location differed from each other [$F(2,30)=145,87,7.5$]. The monaural advantages generally differed significantly from zero [$t(15)>2.9$], except for speech interferers in the −30° and 90° locations. Figure 1 shows that these means were similar to those for the other interferer types and the lack of significance can be attributed to greater variance. Advantage of separation was negative for the interferer at −30°, and positive for 60° and 90°.

The binaural advantage for the interferer at −30° was not calculated since the monaural measurement was not made from the ear with the best signal-to-noise ratio and thus the difference between binaural and monaural measurement includes more than binaural processing in this case. A two-factor ANOVA for the remaining data (4 interferer types×2 interferer locations) revealed no significant effects. The majority of binaural advantages were significantly greater than zero [$t(15)>2.2$] at 60° and 90° and are in the range of 2–4

dB, consistent with previous reports. The exceptions were 90° for speech and modulated noise and 60° for reversed speech.

B. Two interferers

The results for two interferers are shown in Fig. 2.

1. Raw SRTs

The binaural SRTs for two interferers also decrease as the interferers are separated from the target location. However, the ordering of the interferer types is different from that seen in the one-interferer case; the speech interferer now gives among the highest SRTs, while the reversed speech remains the lowest. The relative increase in SRTs against the speech interferers compared to the one-interferer case may reflect an increase in linguistic interference, while the reversed speech retains an advantage due to exploitation of F0 differences. The SRT for the speech interferers is higher than that for the reversed speech interferers by an average of 3.7 dB across locations.

The monaural SRTs were lower than the unseparated condition for the (60°,90°) and (90°,90°) conditions. SRTs in the (−30°,90°) did not differ from the unseparated condition, presumably because the beneficial effect of headshadow is removed when interfering sources are placed on both sides. SRTs for the speech interferer were higher than for the other interferer types. This result contrasts with the single-interferer case, in which speech and reversed speech gave the lowest SRTs.

2. Advantages of separation

The total advantage of separation is up to 12 dB for speech and reversed-speech interferers. Speech and reversed speech had a larger total advantage of separation than modulated noise and noise interferers. This advantage of separation was greater than observed with only a single interferer. The (60°,90°) and (90°,90°) conditions gave a large advantage and the (−30°,90°) a smaller one.

The monaural advantage of separation was subjected to a two-factor ANOVA (4 interferer types×3 interferer locations), which revealed a significant effect of interferer location [$F(2,30)=136$, $p<0.0001$], but not interferer type and no interaction. *Post hoc* comparisons revealed that all levels of interferer location differed [$F(2,30)=238,162,7$]. The monaural advantage for the (−30°,90°) location was not significantly different from zero for the speech and reversed speech interferers [$t(15)<1.6$ in each case], but was significantly *below* zero for the two noise-based interferers [$t(15)>2.3$ in each case]. For all other conditions the monaural advantages were significantly above zero [$t(15)>3$ in each case].

The binaural advantage of separation was subjected to a two-factor ANOVA (4 interferer types×3 interferer locations) of binaural advantages for the two-interferer conditions. The ANOVA revealed a significant effect of interferer type [$F(3,45)=7.1$, $p<0.001$], but no effect of location or interaction. *Post hoc* comparisons of interferer type revealed that speech gave greater binaural advantage than noise and modulated noise [$F(3,45)=17,11$]. All conditions gave mean binaural advantages that were significantly above zero [$t(15)>3.6$ in each case].

The significant effect of interferer type confirms that the origin of the changes in the ordering of the interferer types when a second interferer is introduced result from changes in the effectiveness of binaural processing. With more than one interferer the binaural system is more effective at alleviating interference from a speech or reversed speech source than noise or modulated noise. This effect was replicated in the three-interferer conditions.

C. Three interferers

The results for three interferers are shown in Fig. 3.

1. Raw SRTs

The binaural SRTs decrease as the interferers are separated from the target location. The ordering of the interferer types is similar to that seen for the two-interferer conditions.

The monaural SRTs were lower than the unseparated condition for the (30°,60°,90°) and (90°,90°,90°) interferers, but, as in the two-interferer case monaural SRTs at (−30°,60°,90°), with interferers on both sides, were similar to the unseparated case. SRTs for the speech interferer were higher than for the other interferer types.

2. Advantages of separation

The total advantage of separation is up to 10 dB for speech and reversed speech interferers. As in the two-interferer case, the speech and reversed speech interferers gave a larger total advantage of separation than the two noise-based interferers. Conditions (30°,60°,90°) and (90°,90°,90°) gave a large and similar advantage, while (−30°,60°,90°) gave a smaller advantage.

The monaural advantage was subjected to a two factor ANOVA (4 interferer types×3 interferer locations). The ANOVA revealed a significant effect of interferer location [$F(2,30)=243$, $p<0.0001$] on monaural advantage, but not of interferer type and no interaction. *Post hoc* comparisons of different locations revealed only that the (−30°,60°,90°) condition differed significantly from the (30°,60°,90°) and (90°,90°,90°) conditions [$F(2,30)=153,97$]. Monaural advantage for the (−30°,60°,90°) location was not significantly different from zero for any interferer type, whereas the monaural advantage in all other conditions differed significantly from zero [$t(15)>4.2$ in each case].

A two-factor ANOVA (4 interferer types×3 interferer locations) for binaural advantage revealed a significant effect of interferer type [$F(3,45)=7.7$, $p<0.0005$] and interferer location [$F(2,30)=11.4$, $p<0.0005$], but no interaction. *Post hoc* comparisons of interferer type showed that speech and reversed speech gave consistently larger binaural advantages than did modulated noise or noise interferers [$F(3,45)=13.6,11.3,11.8,9.7$]. Comparisons between interferer locations revealed that binaural advantage in the (90°,90°,90°) condition was significantly different from the other two [$F(2,30)=15.9,18.3$]. Interferer configurations (30°,60°,90°) and (−30°,60°,90°) were not different. How-

ever, binaural advantages in every condition except modulated noise interferers at (30,60,90) were significantly greater than zero [$t(15) > 3.7$].

IV. GENERAL DISCUSSION

The experiment was intended to bridge the gap in complexity from the relatively simple situations that have been extensively researched in previous studies to more complex and realistic listening situations. This was achieved by measuring SRT's both monaurally and binaurally against one, two, or three interferers in four different spatial configurations. In each of these conditions, the interferer was either speech, reversed speech, speech-shaped noise, or speech-modulated noise. The data analysis involved a separation between monaural and binaural effects, making use of the assumption that overall performance is the sum of the effects of best-ear advantage and binaural advantage. The fact that the resulting ''advantage'' measures produce a much simpler and clearer projection of the data than the raw SRTs suggests that this analysis is appropriate. However, the advantages observed for multiple voice-based interferers were larger than can be accounted for by models of binaural unmasking (Zurek, 1992). The patterns of SRTs and spatial advantages revealed a number of effects.

A. Monaural advantage

Monaural listening through the left ear was sufficient to produce an advantage of spatial separation when the interferer(s) all occurred on the right, due to the effect of head-shadow. If one assumes that this advantage arises purely from best-ear listening, the size of this effect is predictable from the acoustics associated with sound waves reaching the head and the importance of the frequencies involved for speech understanding (Zurek, 1992).

The monaural spatial advantage disappeared once multiple interfering sources were spatially distributed on the right and left, since the signal-to-noise ratio for the target presented from front was now reduced by the interferer on the left. Although unsurprising, this effect has important practical implications, since it implies that head-shadow plays a minor role in commonly encountered listening situations when competing sources are distributed in both hemifields. The result also clarifies those of Peissig and Kollmeier (1997). In their study, the fixed sources were always on either side of the head when three interfering sources were used, so their results with three interferers should probably be interpreted as including only effects of binaural advantage.

B. Binaural advantage

When both ears were available to the listener and the target sound was spatially separated from the interferers, a binaural advantage occurred. This advantage has been modeled on the basis of the strength of binaural unmasking at different frequencies and the importance of those frequencies to speech understanding (Levitt and Rabiner, 1967; vom Hövel, 1984; Zurek, 1992). For a single noise interferer, the binaural advantage is predicted to be 3 dB when the spatial separation is 90°. From the present data set, the prediction appears to hold for all interferer types in the one-interferer case. However, for multiple interferers it seems sufficient to explain the data only for noise-based interferers (Fig. 4).

In contrast to the monaural advantage discussed above, the binaural advantage was robust in all spatial configurations, whether competing sources were spatially coincident, distributed across locations, in the same hemifield, or on both the right and left. The role of binaural advantage in complex listening situations is probably greater, therefore, than monaural head-shadow. The fact that binaural advantage was robust against spatially distributed interferers is surprising in the context of models of binaural unmasking that depend upon a highly coherent masker. Multiple interferers with different delays will have reduced coherence and so might be expected to have markedly reduced binaural unmasking. For instance, Durlach's (1963) equalization-cancellation model can cancel an interferer with a specified interaural time delay, but if multiple interferers have multiple delays, one would expect it to be able to cancel only one of them. A follow-up study, Culling *et al.* (2003) analyzes this effect in greater detail and shows that models of binaural unmasking are more robust to reduced coherence than one might expect. On the other hand, it seems unlikely that binaural unmasking can account for all the spatial advantages observed with speech interferers (see Sec. IV E).

C. Dip listening

Another well-known effect is that of ''dip-listening'' where listeners exploit transitory reductions in the power of the interferer in order to pick up information from the target (Festen and Plomp, 1990). Dip listening can be most clearly seen in the current data set through the differences between noise and modulated noise interferers; only the latter gives the listener the opportunity to listen in the dips and thereby achieve a lower SRT. There is a strong effect of dip listening in the single-interferer case of 2–3 dB. As additional interferers are added, the effect is attenuated, because the dips in one interferer become filled in by the energy of another asynchronously modulated interferer (Bronkhorst and Plomp, 1992). In the three-interferer case the SRTs are indistinguishable. Dip listening also, therefore, plays only a minor role in complex listening environments with multiple, relatively distant source like those simulated here.

D. F0 differences

SRTs were lower for single interfering sources that were voiced (speech and reversed-speech) than for ones that were noise-based (noise and modulated noise). The advantage of voiced interferers is seen in the difference between the overall SRTs for these conditions (Fig. 1). In contrast, when two and then three interferers were tested, this difference was not observed. The results may be best understood in terms of a cancellation mechanism that relies on F0 differences (e.g., de Cheveigné 1997), although an informational masking account is also possible (see Sec. IV F). The F0-difference interpretation can account for the fact that the effect is limited to the single-interferer situation, since multiple voices, with

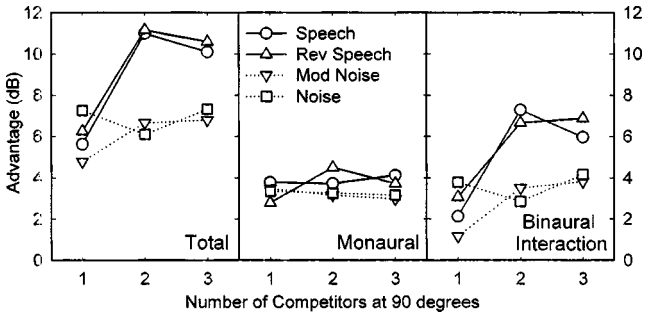


FIG. 4. Spatial advantage as a function of number of interfering sounds at 90° for each interferer type. The left-hand panel shows the total advantage, the middle panel shows the advantage when using only the best ear (monaural advantage), and the right-hand panel shows the difference between these two, attributable to binaural interaction.

multiple F0s, would require multiple rounds of cancellation. It seems likely that the system is incapable of making more than one such cancellation, but even if multiple rounds are possible, the target sound would be progressively distorted by the comb-filtering effects that accompany the cancellation.

SRTs for voiced interferers were substantially higher when there were two or three voiced interferers. The distinction between voiced and noise-based interferers is especially evident in Fig. 5 where the increase in SRT resulting from additional interferers (as large as 14 dB) is compared with the expected increase based on the increased energy in the interferers (3–6 dB). For noise-based interferers, the incremental change in SRT as the second, and then the third, interferers were added can be explained by the increased energy in the interferers (see thick horizontal bars in Fig. 5). In contrast, for the speech and reversed-speech interferers, the incremental change in SRT with added interferers is substantially larger.

E. Voicing/spatial advantage interaction

The interferer type interacted with spatial separation; the effect of spatial separation of interferers from the target was greater when either of the two voiced interferers was used (though only in the two- and three-interferer cases). A similar effect was recently reported by Noble and Perret (2002) and is consistent with the results of Peissig and Kollmeier (1997), who also found that spatial unmasking was more robust with multiple speech interferers than with multiple noise interferers. For the latter binaural advantage is limited to about 3 dB (Bronkhorst and Plomp, 1992). However, the present result is inconsistent with Peissig and Kollmeier's suggested explanation in terms of suppressing different interfering sources at different times. If this explanation were correct, then the speech modulated noise used in the present experiments would also have permitted spatial advantage to be robust against multiple interferers. We have no alternative explanation. However, it is noteworthy that it was a substantial effect (≈3 dB) and was only observed in the most complex and realistic of listening situations. It is therefore worthy of further investigation.

Other than Peissig and Kollmeier's results, the nearest precedents for the effect in the literature are the rather small

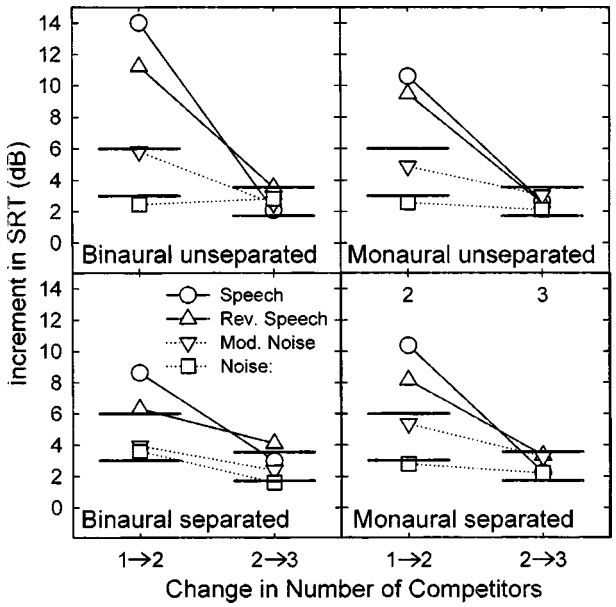


FIG. 5. Change in SRT as each additional interfering sound is added as a function of the total number of interfering sounds. The lower set of horizontal bars shows the expected average increment in threshold when there are random phase relationships between the components of the existing and the added interferers (i.e., there is a 3-dB increase in expected SRT as a result of a 3-dB increase in total masker level when a second interferer is added). The upper set of horizontal bars represent the maximum expected increase in thresholds if the components of the added interferer are perfectly in phase with those of the existing interferers (e.g., adding a second interferer causes a 6-dB increase in total masker level and in SRT).

interactions in "double-vowel" identification reported by Shackleton *et al.* (1994), and later corroborated by Culling *et al.* (1994). Shackleton *et al.* used a design in which the dependent variable was the percentage of simultaneous, synthesized vowel pairs for which listeners correctly identified both vowels. They found an interaction between the presence of a difference in F0 and the presence of a difference in ITD, such that percent correct was higher when two vowels differed in both these parameters. Culling *et al.* used a method more similar to the measurement of SRT in that the threshold for correct vowel identification was measured against a single competing vowel, which varied somewhat from trial to trial. They found a similarly small effect. The effects described in these studies seem to differ in magnitude from the one found here, but the one found here was only evident using multiple interferers. It may be that there is a small interaction for a single interferer and that that interaction grows as more interferers are introduced.

Curiously, the F0 effect, if one defines it as the difference between the reversed-speech and modulated noise conditions, also appears to interact with monaural versus binaural presentation. This interaction may be seen in the one-interferer case (compare the erect and inverted triangles on the top two panels of Fig. 1) where effects of F0 difference are large; they are consistently larger in the binaural than in the monaural condition, regardless of spatial configuration. The reasons for this effect remain obscure.

F. Informational masking

"Informational masking" is disruption to the processing of a target sound without energetically masking it. For in-

stance, the masker may be in different frequency channels or presented to a different ear, and so does not prevent detection of the target. If the content of the interferer is similar to that of the target, the two can become confused and tasks such as target identification can be disrupted (Pollack and Pickett, 1958; Lutfi, 1990; Kidd *et al.*, 1998; Brungart *et al.*, 2001).

One condition for informational masking is that the target content is above detection threshold. In the present study, all the interferers had the same long-term spectrum as the target speech; hence there was always some overlap in the energies of the target and interferers, and energetic masking was always present. Signs of informational masking must, therefore, manifest themselves as an excess masking in particular conditions. In addition, one should expect more informational masking where the overlap in spectro-temporal pattern is relatively incomplete. In the one-interferer cases, the modulated noise contained periods during which the energy in the interferer was significantly reduced. With additional interferers the overlap was more constant for both noise and modulated noise. In contrast, the speech and reversed speech naturally contain dynamic variations in spectrum, and are therefore unlikely to completely overlap in spectrum with the targets at a given instant in time. Thus, one would expect the two voiced interferers to be more likely to display informational masking effects. It is possible that informational masking can be seen in two aspects of the present data set.

First, when multiple interferers were present there was a consistent 2-dB difference between the speech and reversed-speech SRTs. This effect may represent informational masking at the linguistic level, and this suggestion is supported by the fact that when no binaural unmasking is possible (monaural and nonspatially separated configurations) multiple speech interferers produce the highest SRTs of all interferer types. The underlying mechanism is, at this point, largely a matter of speculation. Words from an interfering voice may be intruding into the perceived target sentence. The grammatical and semantic information in the masking stimuli may also be automatically recruiting the listener's attentional resources and reducing the depth of processing that can be applied to the target voice. Using the current paradigm, it is not possible to differentiate the effects of intrusion and attentional distraction; although the listeners transcripts were recorded, the listeners were aware of the content of the interfering sentences, and would have been unlikely to include in their transcripts words that they knew were intrusions. Furthermore, for the two-interferer case, there is evidence that a component of the binaural interaction is a release of this form of informational masking, since there is greater advantage for speech than for reversed speech

Second, the added interference produced by multiple speech and reversed-speech interferers may reflect increased informational masking. This effect was considered above with respect to the effect of F0 differences, but an increase in informational masking may provide an alternative explanation. This account relies upon the reversed speech acting as an informational masker at a lower linguistic level. It is possible, for instance, that reversed speech can recruit attentional and cognitive resources that noise-based interferers do not because they engage phonetic and lexical processing resources even if they do not yield meaningful lexical units for higher levels of processing. It is possible that reversed-speech engages many of these processes by activating an initial mechanism that searches for sources containing language-based information.

Informational masking perhaps offers a more coherent account of the interaction between voicing of the interferers and spatial separation because informational masking can be released by the spatial separation (Brungart *et al.*, 2001). However, there are other problems with this account. First, the effect of F0 differences is very well established and reduction of this effect must account for at least some of the increase in SRTs that occurs as a second voice-based interferer is introduced. Second, in the multiple-interferer cases, thresholds in the reversed speech condition are no worse than for the two noise-based interferers. Thus, there is no obvious evidence of an additional masking effect for the reversed speech interferer with respect to other interferer types, only with respect to the single-interferer case.

It should be possible to differentiate between aspects of the current data set that can be explained by informational masking and those that can be explained by F0 differences by repeating elements of the experiment using an additional masker type. Shannon *et al.* (1995) showed that very accurate speech recognition could be achieved by listening to a noise that was modulated within a discrete number of frequency channels by the speech envelope within those channels. If a sufficiently small number of frequency channels is used, such speech lacks an F0, but should still possess many of the attributes necessary to cause both types of informational masking considered above. If such interferers show a pattern of thresholds similar to the speech interferers in the current study, then this finding would strongly support the informational masking account.

Although the effects that can unambiguously be attributed to informational masking in the current data set are not very large, it should be noted that some aspects of the SRT paradigm we employed were not optimal for the observation of informational masking. In particular, the use of a fixed interfering sentence or set of sentences throughout a given SRT measurement and the presentation of the text of the interfering messages at the beginning of the measurement will have substantially reduced the uncertainty about the interferer content. Uncertainty about the interferer is supposed to be a vital aspect of informational masking, so this methodology may have served to reduce the size of the effects observed.

V. CONCLUSIONS

The results obtained in the present study suggest that listeners' ability to function in complex environments, such as a cocktail party, not only depends the type, number, and location of interfering sounds, but also on interactions between these factors. A number of the effects observed in the current study are well established, but the interactions between interferer types and spatial configuration have not been previously reported and not always easily explained. Further research is necessary to explore and account for these phenomena. However, from a practical point of view

the most significant finding is that in complex listening environments the effects of binaural advantage and fundamental frequency difference seem to be interdependent, while the role of dip listening is reduced. These findings both clarify our understanding of the cocktail party problem and its solution, and should inform our choice of appropriate stimuli for clinical testing of binaural processing.

ACKNOWLEDGMENTS

This work was supported by UK MRC and NIH-NIDCD Grant Nos. R01-DC00100 and R29-DC03083.

[1]The recordings were obtained from Patrick Zurek, Ph.D., of the Massachusetts Institute of Technology.

[2]The SRT was the mean of the levels on trials 4, through to 11, inclusive. Although no 11th trial existed, the level at which it would have been presented was determined by the result of the tenth trial.

[3]Subjects from the two conditions were paired according to the order in which they took part; subject 1 from the monaural condition was paired with subject 1 from the binaural condition. These pairs experienced the within-subjects conditions and materials in the same order but using a different number of ears. Alternative pairings would have no effect on the means of the resulting differences, but would have resulted in larger variances due to variations in list difficulty.

Blauert, J., Brueggen, M., Bronkhorst, A. W., Drullman, R., Reynaud, G., Pellieux, L., Krebber, W., and Sottek, R. (**1998**). ''The AUDIS catalog of human HRTFs,'' J. Acoust. Soc. Am. **103**, 3082.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). ''Binaural processing model based on contralateral inhibition. I. Model structure,'' J. Acoust. Soc. Am. **110**, 1074–1089.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001b**). ''Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters,'' J. Acoust. Soc. Am. **110**, 1089–1105.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001c**). ''Binaural processing model based on contralateral inhibition. III. Dependence on temporal parameters,'' J. Acoust. Soc. Am. **110**, 1105–1118.

Bregman, A. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge).

Bronkhorst, A. (**2000**). ''The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions,'' Acustica **86**, 117–128.

Bronkhorst, A. W., and Plomp, R. (**1988**). ''The effect of head-induced interaural time and level differences on speech intelligibility in noise,'' J. Acoust. Soc. Am. **83**, 1508–1516.

Bronkhorst, A. W., and Plomp, R. (**1992**). ''Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing,'' J. Acoust. Soc. Am. **92**, 3132–3139.

Brokx, J. P. L., and Nooteboom, S. G. (**1982**). ''Intonation and the perceptual separation of simultaneous voices'' J. Phonetics **10**, 23–36.

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (**2001**). ''Informational and energetic masking effects in the perception of two simultaneous talkers,'' J. Acoust. Soc. Am. **109**, 1101–1109.

Cherry, E. C. (**1953**). ''Some experiments on the recognition of speech with one and two ears,'' J. Acoust. Soc. Am. **25**, 975–979.

Culling, J. F., and Summerfield, Q. (**1995**). ''Perceptual segregation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay,'' J. Acoust. Soc. Am. **98**, 785–797.

Culling, J. F., Summerfield, Q., and Marshall, D. H. (**1994**). ''Effects of simulated reverberation on binaural cues and fundamental frequency differences for separating concurrent vowels,'' Speech Commun. **14**, 71–96.

Culling, J. F., Hawley, M. L., and Litovsky, R. Y. (**2003**). ''The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources,'' submitted to J. Acoust. Soc. Am. .

Darwin, C. J., and Culling, J. F. (**1990**). ''Speech perception seen through the ear,'' Speech Commun. **9**, 469–476.

de Cheveigné, A. (**1997**). ''Concurrent vowel identification. III. A neural model of harmonic interference cancellation,'' J. Acoust. Soc. Am. **101**, 2857–2865.

de Cheveigné, A., McAdams, S., Laroche, J., and Rosenberg, M. (**1995**). ''Identification of concurrent harmonic and inharmonic vowels: A test of the theory of harmonic cancellation,'' J. Acoust. Soc. Am. **97**, 3736–3748.

Durlach, N. I. (**1963**). ''Equalization and cancellation model of binaural unmasking-level differences,'' J. Acoust. Soc. Am. **35**, 1206–1218.

Festen, J. M. (**1993**). ''Contributions of comodulation masking release and temporal resolution to the speech-reception threshold masked by an interfering voice,'' J. Acoust. Soc. Am. **94**, 1295–1300.

Festen, J. M., and Plomp, R. (**1990**). ''Effects of fluctuating noise and interfering speech on the speech-reception SRT for impaired and normal hearing,'' J. Acoust. Soc. Am. **88**, 1725–1736.

Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (**1999**). ''Speech intelligibility and localization in complex environments,'' J. Acoust. Soc. Am. **105**, 3436–3448.

Kidd, Jr., G., Mason, C. R., Rohtla, T. L., and Deliwala, P. S. (**1998**). ''Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns,'' J. Acoust. Soc. Am. **104**, 422–431.

Koehnke, J., and Besing, J. M. (**1996**). ''A procedure note for testing speech intelligibility in a virtual listening environment,'' Ear Hear. **17**, 211–217.

Lea, A. (**1992**). ''Auditory models of vowel perception,'' Ph.D. thesis, Nottingham (unpublished).

Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Levitt, H., and Rabiner, L. R. (**1967**). ''Predicting binaural gain in intelligibility and release from masking for speech,'' J. Acoust. Soc. Am. **42**, 620–629.

Litovsky, R. Y., Fligor, B., and Tramo, M. (**2002**). ''Functional role of the human inferior colliculus in binaural hearing,'' Hear. Res. **165**, 177–188.

Lutfi, R. A. (**1990**). ''How much masking is informational masking?'' J. Acoust. Soc. Am. **88**, 2607–2610.

MacKeith, N. W., and Coles, R. R. A. (**1971**). ''Binaural advantages in hearing speech,'' J. Laryngol. Otol. **85**, 213–232.

Noble, W., and Perret, S. (**2002**). ''Hearing speech against spatially separate competing speech versus competing noise,'' Percept. Psychophys. **64**, 1325–1336.

Nilsson, M., Soli, S. D., and Sullivan, J. A. (**1994**). ''Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise,'' J. Acoust. Soc. Am. **95**, 1085–1099.

Peissig, J., and Kollmeier, B. (**1997**). ''Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners,'' J. Acoust. Soc. Am. **101**, 1660–1670.

Plomp, R. (**1986**). ''A signal-to-noise ratio method for the speech-reception SRT of the hearing impaired,'' J. Speech Hear. Res. **29**, 146–154.

Plomp, R., and Mimpen, A. M. (**1979**). ''Speech-reception threshold for sentences as a function of age and noise level,'' J. Acoust. Soc. Am. **66**, 1333–1342.

Plomp, R., and Mimpen, A. M. (**1981**). ''Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech-reception threshold for sentences,'' Acustica **48**, 325–328.

Pollack, I., and Pickett, J. M. (**1958**). ''Stereophonic listening and speech intelligibility against voice babbles,'' J. Acoust. Soc. Am. **30**, 131–133.

Rothauser, E. H., Chapman, W. D., Guttman, N., Nordby, K. S., Silbiger, H. R., Urbanek, G. E., and Weinstock, M. (**1969**). ''I.E.E.E. recommended practice for speech quality measurements,'' IEEE Trans. Audio Electroacoust. **17**, 227–246.

Shackleton, T. M., Meddis, R., and Hewitt, M. J. (**1994**). ''The role of binaural and fundamental frequency differences in the identification of concurrently presented vowels,'' Q. J. Exp. Psychol. A **47A**, 545–563.

Shannon, R. V., Zeng, F.-G., Kamath, J., Wygonski, J., and Ekelid, M. (**1995**). ''Speech Recognition with Primarily Temporal Cues,'' Science **270**, 303–304.

Shinn-Cunningham, B. G., Schickler, J., Kopco, N., and Litovsky, R. Y. (**2001**). ''Spatial unmasking of nearby speech sources in a simulated anechoic environment,'' J. Acoust. Soc. Am. **110**, 1118–1129.

Summerfield, Q. (**1992**). ''Roles of harmonicity and coherent frequency modulation in auditory grouping,'' in *The Auditory Processing of Speech*, edited by M. E. H. Schouten (Mouton de Gruyter, New York).

Summerfield, Q., and Culling, J. F. (**1992**). ''Periodicity of maskers not targets determines ease of perceptual segregation using differences in fundamental frequency,'' J. Acoust. Soc. Am. **92**, 2317(A).

vom Hövel, H. (**1984**). ''Zur Bedeutung der bertragungseigenschaften des Außenohres sowie des binauralen Hörsystems bei gestörter Sprachübertragung,'' Ing. Dissertation, RWTH Aachen 1984; as cited by Bronkhorst, A. (**2000**). Acustica **86**, 117–128.

Zurek, P. M. (**1992**). ''Binaural advantages and directional effects in speech intelligibility,'' in *Acoustical Factors Affecting Hearing Aid Performance*, edited by G. A. Studebaker and I. Hochberg (Allyn and Bacon, Boston), pp. 255–276.

Better place-coding of the fundamental frequency in cochlear implants

Luc Geurts and Jan Wouters
Laboratory for Experimental ORL, KULeuven, Kapucijnenvoer 33, B 3000 Leuven, Belgium

(Received 31 October 2002; revised 3 November 2003; accepted 24 November 2003)

In current cochlear implant systems, the fundamental frequency $F0$ of a complex sound is encoded by temporal fluctuations in the envelope of the electrical signals presented on the electrodes. In normal hearing, the lower harmonics of a complex sound are resolved, in contrast with a cochlear implant system. In the present study, it is investigated whether ''place-coding'' of the first harmonic improves the ability of an implantee to discriminate complex sounds with different fundamental frequencies. Therefore, a new filter bank was constructed, for which the first harmonic is always resolved in two adjacent filters, and the balance between both filter outputs is directly related to the frequency of the first harmonic. The new filter bank was compared with a filter bank that is typically used in clinical processors, both with and without the presence of temporal cues in the stimuli. Four users of the LAURA cochlear implant participated in a pitch discrimination task to determine detection thresholds for $F0$ differences. The results show that these thresholds decrease noticeably for the new filter bank, if no temporal cues are present in the stimuli. If temporal cues are included, the differences between the results for both filter banks become smaller, but a clear advantage is still observed for the new filter bank. This demonstrates the feasibility of using place-coding for the fundamental frequency. [DOI: 10.1121/1.1642623]

PACS numbers: 43.66.Ts, 43.66.Fe, 43.66.Hg [MRL] Pages: 844–852

I. INTRODUCTION

One of the first steps in sound processors for cochlear implants is a splitting of the audio signal into several frequency bands or channels, covering the speech frequency range. Since these bands are relatively broad, individual harmonics of periodic sounds are unlikely to be resolved, resulting in beats at the output of the filters. These temporal fluctuations in the envelope provide information on the fundamental frequency $F0$ of the sound. If these fluctuations are still present in the electrical signals that stimulate the auditory nerve, cochlear implantees can use this source of information to detect $F0$ differences. Sinusoidally amplitude-modulated pulse trains contain similar fluctuations and the elicited pitch increases with modulation frequency. Some implantees can detect differences in the modulation frequency as small as 2%, the exact value being dependent on the modulation frequency and the subject (McKay *et al.*, 1994). Geurts and Wouters (2001) used synthetic vowels to measure the detectable $F0$ difference, and found values between 4% and 13%. In general, the just-noticeable frequency difference increases with frequency, whether the stimuli are pulse trains with varying rate (Eddington *et al.*, 1978; Shannon, 1983; Tong and Clark, 1985; Townshend *et al.*, 1987), amplitude-modulated pulse trains (McKay *et al.*, 1994), or synthetic vowels, processed by a sound processor (Geurts and Wouters, 2001). The frequency limit above which rate discrimination thresholds are very large or even unmeasurable lies on the average around 300 Hz (Eddington *et al.*, 1978; Townshend *et al.*, 1987; McKay *et al.*, 1994). Although this limit is above common values of $F0$ for speech sounds, it is far below the upper limit of common $F0$'s in music. Pijl and Schwarz (1995) showed that melody recognition is possible with single-electrode pulse trains, but performance dropped when the values of the $F0$'s in the melody increased. They concluded that temporal cues only are sufficient for the mediation of musical pitch with a cochlear implant, but mainly for the lower half of the range of $F0$'s commonly used in music, i.e., up to about 500 Hz.

In contrast with a cochlear implant processor, the lowest harmonics are resolved in acoustic normal hearing, and the places of maximal activity along the basilar membrane might also provide $F0$ information. It is well known that the lower-resolved harmonics dominate pitch perception. The dominance of each harmonic depends on the frequency of the fundamental and on the subject (Moore, 1997). Given the broad filters used in a cochlear implant processor, all the harmonics, even the lowest, are unlikely to be resolved. This difference with acoustic normal hearing might contribute to the worse performance of cochlear implantees in pitch-discrimination tasks.

In this article, a new filter bank will be suggested, which attempts to improve the place coding of an individual harmonic. The idea is based on a finding of McDermott and McKay (1994), stating that two concurrent pulse trains on two adjacent electrodes elicit a single pitch, its value depending on the relative difference between the levels of the pulses on both channels. If both pulse trains sound equally loud when presented separately, then the combined stimulation will yield a pitch lying in between the pitches of the pulse trains on the separate channels. The new filter bank is designed in such a way that when a single-frequency component moves through two adjacent filters, the output of the ''lower'' filter will gradually drop and the output of the ''higher'' filter will gradually rise. This filter bank will be compared in a pitch-discrimination task with ''classic'' filters, having broad, flat passbands, and steep slopes beyond

the cutoff frequencies. This comparison permits an evaluation of whether the induction of resolved harmonics improves pitch perception of complex stimuli. Also, the importance of the temporal cue will be evaluated, by comparing algorithms that induce temporal $F0$ fluctuations in the output with algorithms yielding "flat" output signals.

II. SOUND PROCESSING

In total, four algorithms based on the Continuous Interleaved Sampling (CIS) strategy (Wilson *et al.*, 1991) are used in the perception experiments. Either the classic or the new filter bank is used, and temporal $F0$ fluctuations are allowed or not. Each algorithm consists of the following steps: bandpass filtering, envelope detection, compression, and pulse-train modulation.

A. Bandpass filters

For the classic filter bank, the input signal is sampled at 10 kHz, as is the case in the clinical LAURA speech processor. The filter bank consists of seven or eight bandpass filters, the number varying among participating subjects. The filters are eighth-order Butterworth bandpass filters with cutoff frequencies of 100, 325, 550, 775, 1125, 1634, 2372, 3444, and 4999 Hz in the eight-channel case and 100, 357, 614, 934, 1421, 2161, 3287, and 4999 Hz in the seven-channel case. These values correspond to the settings of the clinical LAURA processors of the participants. Approximately, the filters are equally spaced on a linear scale below 1 kHz and on a logarithmic scale above. At low frequencies, the filters are relatively broad, which is characteristic for all currently available cochlear implant systems.

For the new filter bank, referred to as the triangle filter bank, the input signal is sampled at 16 kHz, and at later stages in the process further subsampled to 8 kHz, 4 kHz, 2 kHz, 1 kHz, and 500 Hz (see below). This filter bank is designed to resolve the first harmonic of a complex stimulus in two adjacent filters. Ideally, the width of each filter is maximum one octave, meaning that the filter response should be zero at and below the lowest boundary frequency, and at and above twice this frequency. The maximum response is obtained half an octave above or below each boundary frequency. This maximum will be set to 1 (0 dB).

The filter is based on a simple loudness model for normal hearing. It is designed such that the loudness of a pure tone sweeping through the filter increases linearly with frequency from the lower boundary frequency f_{BL} to the center frequency f_C, and decreases linearly from the center frequency to the upper boundary frequency f_{BH}

$$L = k_1 \cdot (f - f_{BL}) \quad \text{for } f_{BL} \leqslant f \leqslant f_C, \tag{1}$$

$$L = k_1 \cdot (f_{BH} - f) \quad \text{for } f_C \leqslant f \leqslant f_{BH}. \tag{2}$$

According to Stevens' law, loudness is proportional to intensity to the power 0.3 (Stevens, 1957)

$$L = k_2 \cdot I^{0.3}. \tag{3}$$

This implies that, for a normal-hearing subject, the intensity of a pure tone passed through the filter should be proportional to the frequency difference $f - f_{BL}$ (or $f - f_{BH}$) to the power 10/3

$$I = k_3 \cdot (f - f_{BL})^{10/3} \quad \text{for } f_{BL} \leqslant f \leqslant f_C, \tag{4}$$

$$I = k_3 \cdot (f_{BH} - f)^{10/3} \quad \text{for } f_C \leqslant f \leqslant f_{BH}. \tag{5}$$

Since intensity is the square of amplitude, the gain in the passband should be proportional to the frequency difference to the power 5/3. The constant of proportionality is chosen such that the gain is 1 at the center frequency. This yields

$$\text{gain} = \left(\frac{f - f_{BL}}{f_C - f_{BL}} \right)^{5/3} \quad \text{for } f_{BL} \leqslant f \leqslant f_C, \tag{6}$$

$$\text{gain} = \left(\frac{f - f_{BH}}{f_C - f_{BH}} \right)^{5/3} \quad \text{for } f_C \leqslant f \leqslant f_{BH}. \tag{7}$$

Note that this formula is also applicable to cochlear implantees. At this stage the processing is still linear, and the filter simply adjusts the amplitude of an acoustical stimulus, such that its loudness varies linearly with frequency for a normal-hearing subject. In a later stage, the amplitude will be nonlinearly mapped to a current, using a compression function that takes the loudness growth of electrical hearing into account. In the ideal case of a perfect loudness growth function, the loudness will also vary linearly with frequency for a cochlear implantee.

Adjacent filters overlap considerably. The center frequency of a filter corresponds to the higher boundary frequency of the adjacent filter at the low-frequency side, and to the lower boundary frequency of the adjacent filter at the high-frequency side. Figure 1 gives the desired response of two adjacent filters (circles or squares) and their approximation (solid lines). The horizontal axis denotes frequency relative to half the sampling rate. The center frequencies of the filters are 0.44 and 0.62 times half the sampling rate. Since loudness varies linearly with frequency according to the model, a frequency increase should result in a loudness drop

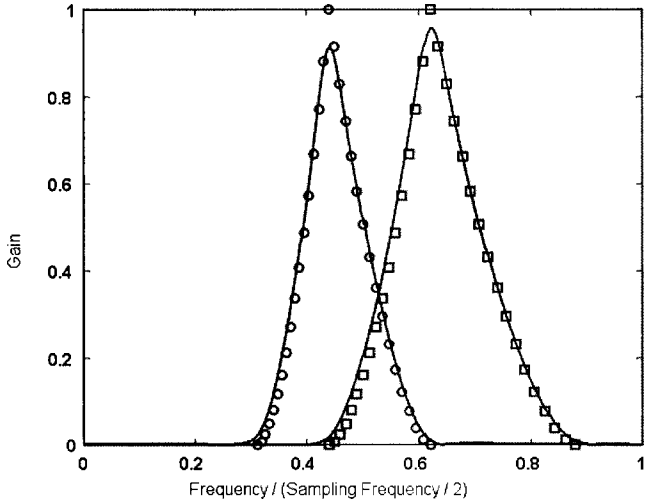


FIG. 1. The frequency characteristics of two adjacent triangular filters. The horizontal axis denotes frequency relative to half the sampling frequency. The circles and the squares indicate the desired response. The solid lines show the response of the eventually used implementation of the filters. For the lowest sampling frequency of 500 Hz used in the experiments, the center frequencies correspond to 110 and 155 Hz.

in the lower filter and an equal and opposite loudness growth in the higher filter. The ideal response is approximated with a 16th-order infinite impulse response (IIR) filter, using the MATLAB[1] implementation of a modified Yule–Walker routine (Friedlander and Porat, 1984). The true filter gain is −33 and −40 dB at the lower and higher boundary frequency, respectively. In the stop band, defined as the region beyond the zeros (−∞ dB) of the filter, the attenuation is at least 48 dB.

The full filter bank is implemented using a so-called tree structure. The outputs of the two highest filters (Fig. 1) are calculated, at a sample rate of 16 kHz. The input signal also goes through a third ''antialiasing'' filter (seventh-order elliptic low-pass filter) before it is sampled down by a factor of 2. Next, the combined process of three filter operations is repeated another five times. In the last stage, the lowest sampling frequency is 500 Hz, and the center frequencies of the corresponding triangular filters are 110 and 155 Hz. In this way, 12 filters are created. Unfortunately, the LAURA participants only have seven or eight available channels. Therefore, several pairs of the highest filters are combined to form one filter with a flat passband between the center frequencies of the individual filters. This implies that the intended place coding of frequency will not work properly at those higher frequencies. Yet, the principle can easily be extended to higher frequencies if enough channels are available. Figure 2 shows the frequency characteristics of both filter banks (classic and triangle) in the eight-channel case. Note the broad, flat passbands in the classic filter bank.

B. Envelope detection, compression, and pulse modulation

In the next stage, the envelopes of the filter outputs are derived using a rectifier and a fourth-order low-pass smoothing filter. To obtain an envelope without $F0$ fluctuations, a full-wave rectifier is used, and the cutoff frequency of the smoothing filter is set to 20 Hz. This will be referred to as the flat condition. Clear temporal $F0$ fluctuations in the envelope are obtained if a half-wave rectifier is used in combination with a cutoff frequency of 250 Hz. This will be referred to as the temp condition, and results in modulation depths at which maximum pitch discrimination scores are obtained (Geurts and Wouters, 2001). The envelope is compressed using a logarithmic function, such that an envelope value of zero is mapped to the current threshold level of the corresponding electrode channel, and the maximum envelope to the most comfortable level. These levels are obtained for each channel before the experiments start. According to the CIS strategy, all envelopes are modulated with interleaved pulse trains. The biphasic pulses have a pulse width of 40 μs per phase, and are stimulated on bipolar channels. The pulse rate is either 1250 pps (8 channels) or 1429 pps (7 channels), yielding an overall pulse rate of 10 000 pps.

III. EXPERIMENT

A. Subjects

Four postlingually deafened users of the LAURA[2] cochlear implant participated in this study. Some details about the subjects can be found in Table I. Two versions of the electrode array of the LAURA implant exist, and both consist of eight bipolar electrode pairs, called channels. These are numbered from apex to base. Subjects SV, JH, and DD use type 5.6: the distance between the two electrodes of one channel is 1.3 mm, and the distance between the midpoints

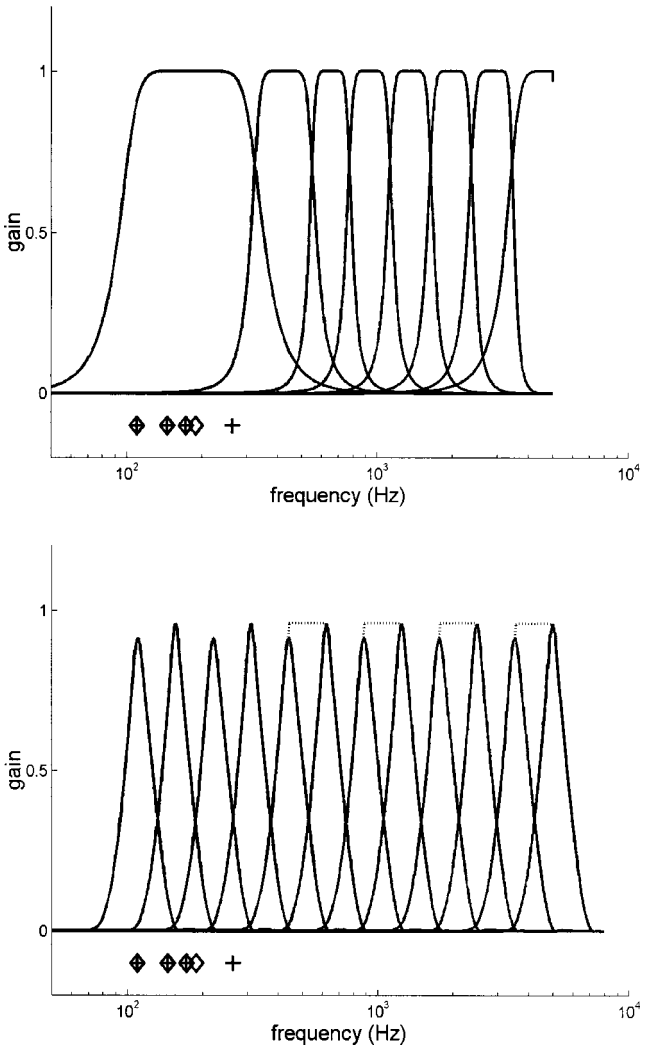


FIG. 2. The frequency characteristics of the classic filter bank (above) and the new triangle filter bank (below) in the case that eight electrode channels are available. The classic filter bank contains somewhat broader filters. The four highest pairs of triangular filters are combined to four single filters with a flat passband between the center frequencies of the individual filters. In the seven-channel case, also filters three and four are combined similarly. The symbols below the horizontal axis indicate the values of the fundamental frequencies of the standard stimuli of the different conditions: diamonds for seven-channel stimuli, plus-signs for eight-channel stimuli.

TABLE I. Some details of the four LAURA subjects who participated in this study. The same channels are used daily in the subjects' clinical processors.

Subject	Age (yrs)	Duration of profound deafness (yrs)	Etiology	Implant experience (yrs)	Number of active electrode channels
DD	36	1	Trauma	6	7
WM	51	5	Progressive	4	7
JH	20	3	Meningitis	7	7
SV	33	6	Unknown	7	8

TABLE II. Formant frequencies (Fi) and bandwidths (Bi) in Hz of the synthetic vowels used in experiment II (Govaerts, 1978).

Vowel	$F1$	$B1$	$F2$	$B2$	$F3$	$B3$	$F4$	$B4$	$F5$	$B5$
/a/	890	110	1360	120	2475	160	3530	275	4830	325
/i/	225	40	2280	90	2820	120	3200	120	3500	150

of adjacent channels is 2.05 mm. For subject WM, who uses type 5.8, these values are, respectively, 2 and 3 mm, which means that his channels physically overlap. Both array types are schematically presented in Fig. 1 of van Wieringen and Wouters (1999).

B. Stimuli

All stimuli were synthetic vowels, created using the Klatt synthesizer available in the CSRE software package.[3] The duration of the stimuli was 500 ms, and a sample rate of 20 kHz was chosen. The formant frequencies and bandwidths corresponded either to /a/ or /i/, and were kept constant for all fundamental frequencies. The exact values are typical for Flemish/Dutch spoken language and are given in Table II (Govaerts, 1978). When using the Klatt synthesizer, the value of the fundamental period has to equal an integer multiple of the sampling period. The fundamental period of the synthesized vowels ranged from 2.7 ms ($F0 = 370$ Hz) to 10.0 ms ($F0 = 100$ Hz).

After the synthesis, the stimuli were further processed using custom-made MATLAB functions. The first 100 ms were truncated to excise the nonzero dc level at the onset. Second, the vowels were gated at the initial 50 ms and at the final 50 ms, using a cosine ramp. Finally, all vowels were balanced in root-mean-square level. Note that this does not guarantee that loudness cannot be used as a cue when two vowels with different $F0$ must be discriminated.

Eventually, these stimuli were processed by the speech-processing algorithms. For this study, it is of interest to know how the energy in each channel varies with the fundamental frequency of the vowel. Figure 3 shows the average amplitude in each of the first three channels as a function of $F0$, obtained after bandpass filtering and envelope detection, but before compression (see supra). Note that this amplitude corresponds to the steady-state value in the flat condition, and that this value is proportional to the mean value of the fluctuating envelope in the temp condition. The upper and the lower plots correspond to the results for the vowel /i/ and /a/, respectively. Note the typical triangular shape of the curves for the triangle filter bank (right plots). It is clear that the channel outputs could provide a consistent source of information on the value of $F0$. This is not the case for the classic filter banks, where the energy either hardly changes with $F0$ or fluctuates in an inconsistent and unpredictable way, as long as the fundamental component is within the passband. From this, it is predicted that it is more difficult to derive a consistent cue on the value of $F0$.

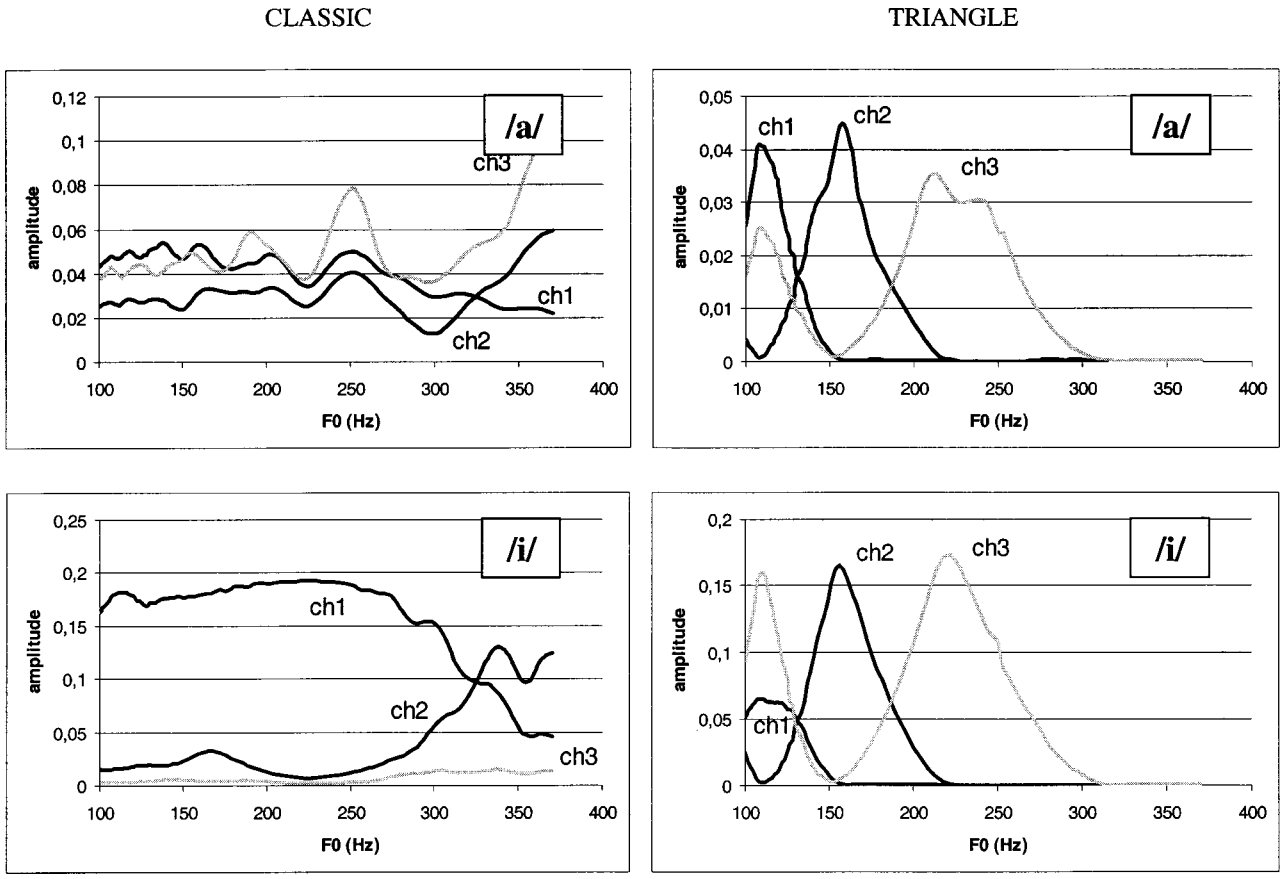


FIG. 3. The average amplitude in the first three channels (out of eight) as a function of the fundamental frequency for the vowels /a/ (upper plots) and /i/ (lower plots). The left plots are the results for the classic filter bank, the right plots for the triangle filter bank.

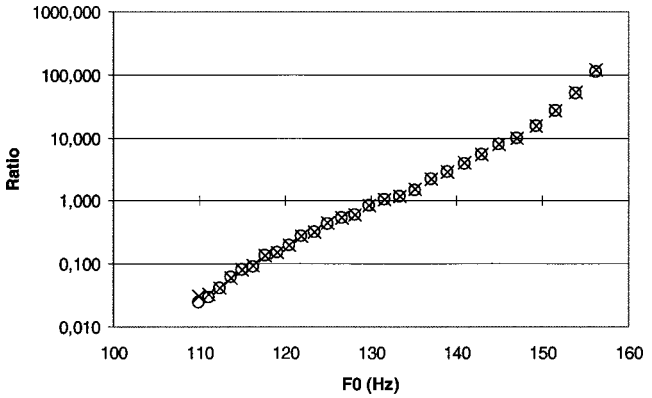


FIG. 4. The ratio of the average uncompressed amplitudes of channel 1 and 2 as a function of $F0$ between the center frequencies of both channels using the triangle filter bank. The circles indicate the values for the vowel /a/, the crosses for the vowel /i/.

The data for the triangle filter bank are further analyzed by calculating the ratio of the average uncompressed amplitudes of channel 1 and 2 as a function of $F0$ between the center frequencies of both channels. Since this ratio only depends on the frequency of the fundamental component, and not on its amplitude, it is almost perfectly independent of spectral shape, as shown in Fig. 4. So, the new coding strategy seems to provide a consistent physical cue, which might be used by implantees.

C. Method

To estimate the smallest difference in $F0$ that could be discriminated, a 2-down, 1-up adaptive procedure was used (Levitt, 1971). Two vowels with different $F0$, further called the *standard* and the *signal*, were presented in random order, and the subject had to indicate the stimulus with the higher pitch.[4] Correct response feedback was not given. $F0$ of the standard stimulus was fixed to either 110, 145, 172, or 189 Hz for the subjects using seven channels. The latter value was changed to 263 Hz for the subject using eight channels. These frequencies are indicated on the horizontal axis of Fig. 2. They were chosen such that different segments on the triangular shape were ''visited'' during the tests: at the top, at the crossings of neighboring filters, and in between the crossings. Note that all these values fall within the first filter of the classic filter bank. Hence, there were 32 conditions in total: two phonemes (/a/ and /i/), four standard $F0$'s, and four algorithms,[5] and each condition was tested once.

For the first trial, the $F0$ of the signal was at least 40% above the $F0$ of the standard, except for the highest $F0$ of 263 Hz for which the signal's $F0$ started at 333 Hz (+27%).

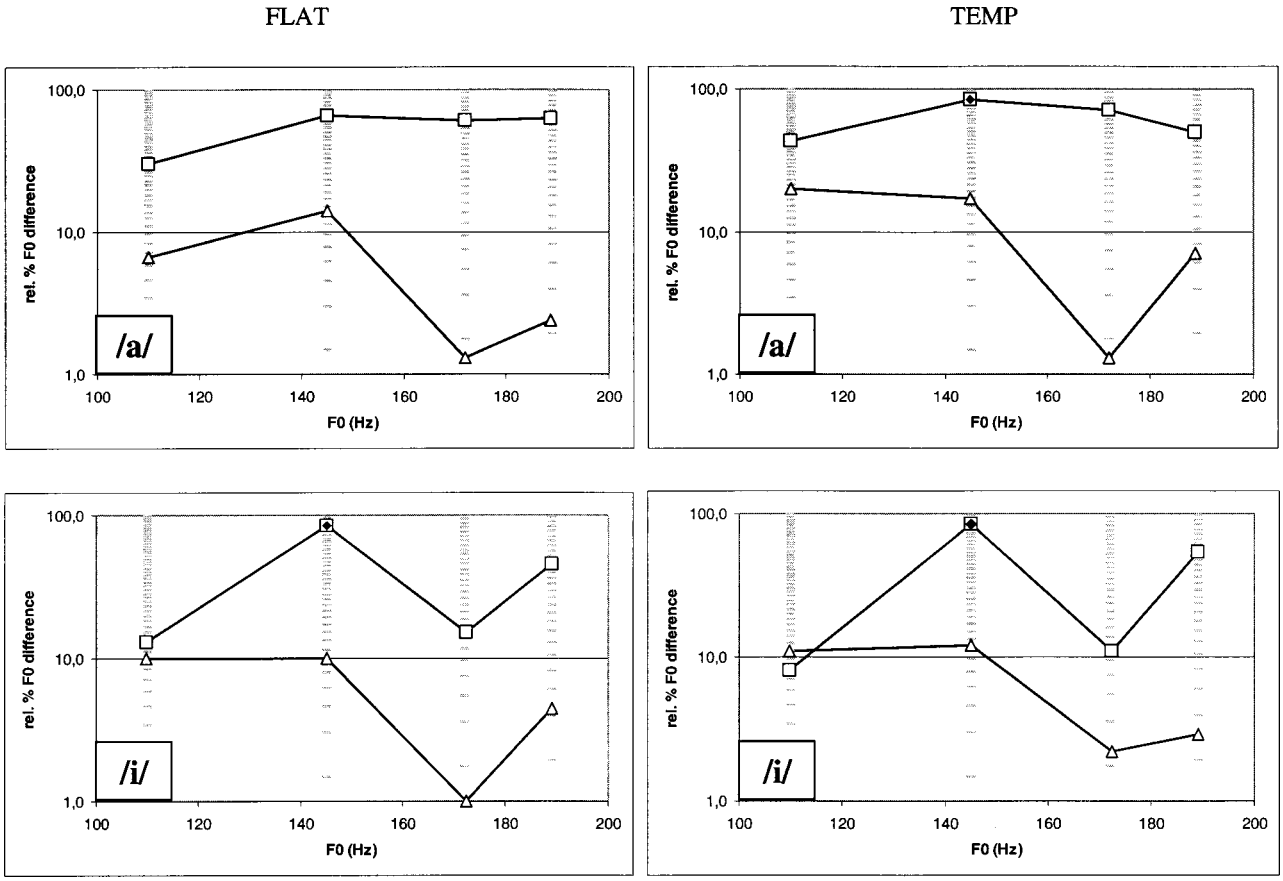


FIG. 5. The just-noticeable $F0$ difference (jnd) for different values of the standard $F0$ for subject DD, using the classic filter bank (squares) or the new filter bank (triangles). The left plots show the results for the stimuli without temporal cues, the right plots those for the stimuli with temporal cues. The synthetic vowel was /a/ or /i/ (respectively, upper and lower plots). A diamond in a symbol indicates that the adaptive procedure was diverging. Since there was a limit on the largest $F0$ difference presented, the obtained value is an underestimate of the real noticeable $F0$ difference. The small gray horizontal lines indicate the relative $F0$ values of all stimuli that could be visited in the adaptive procedure. For clarity, the lines connect points within the same condition and are not intended to represent a relationship between $F0$ and the jnd.

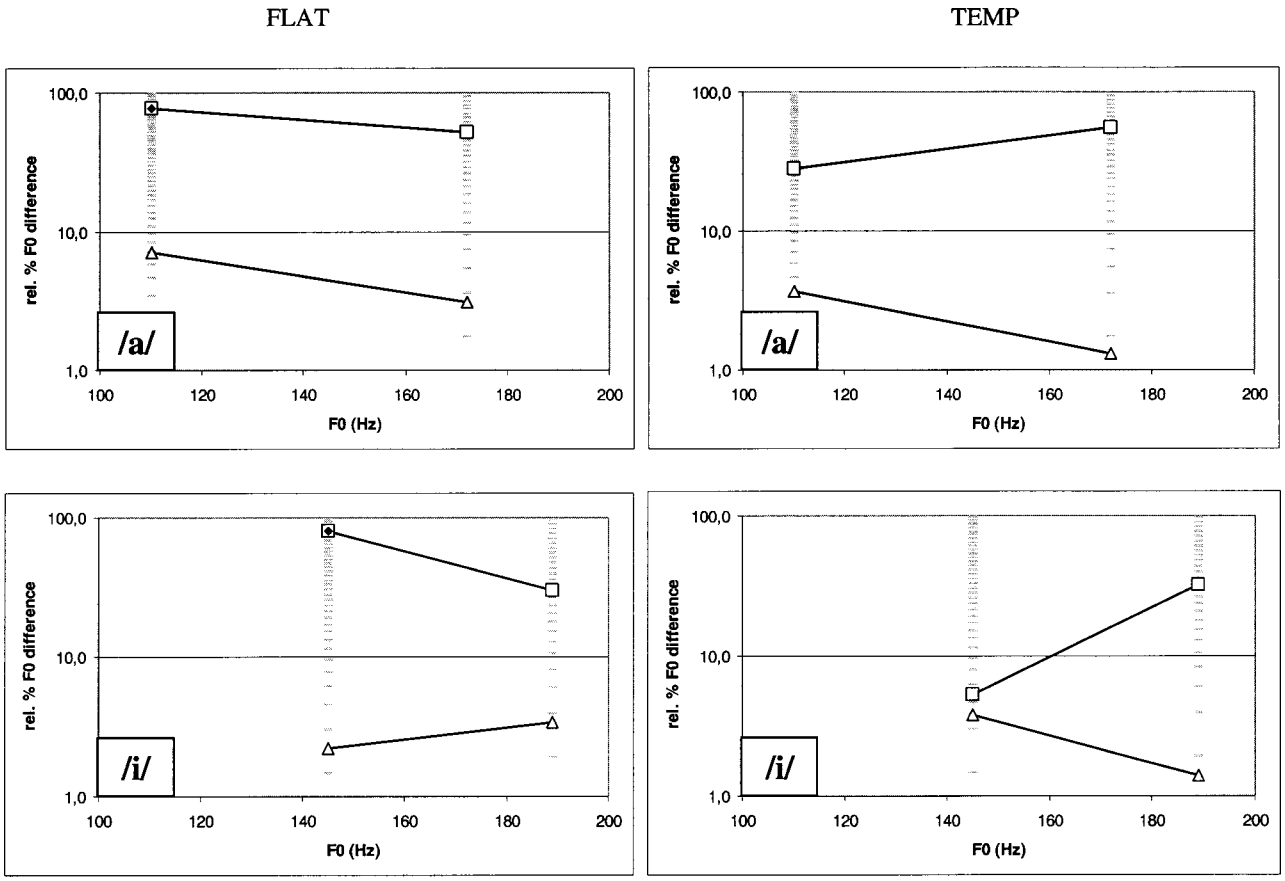


FIG. 6. Same as Fig. 5 for subject WM.

After two successive correct answers, the $F0$ of the signal was lowered, while after one incorrect response, it was increased. Initially, the step size of the fundamental period was 0.2 ms (standard $F0 = 263$ Hz) or 0.4 ms (other $F0$'s). After two reversals, the step size was reduced to 0.05 or 0.1 ms, respectively. The procedure stopped after ten reversals, and the mean value of the fundamental period at the last four reversals was calculated. This value corresponds to the fundamental period of a signal that could be discriminated in approximately 71% of the cases (Levitt, 1971). For the programming and the execution of the experiments, including the presentation of the stimuli, the APEX setup was used (Geurts and Wouters, 2000).

D. Results

Figures 5–8 show for each subject the just-noticeable $F0$ differences (jnd) for the different conditions. Squares indicate the results for the classic filter bank, triangles those for the new filter bank. The plots on the left of each figure show the results for the stimuli without temporal cues (flat), the right plots those for the stimuli with temporal cues (temp). The results for the vowels /a/ and /i/ are shown in the upper and lower plots, respectively. A diamond in a symbol indicates that the adaptive procedure was diverging, i.e., presented $\Delta F0$'s grew. Since there was a limit on the largest $F0$ difference presented, the obtained value is an underestimate of the real noticeable $F0$ difference.

Subject DD (Fig. 5) obtained markedly lower thresholds using the triangle filter bank in all but one condition (exception: vowel /i/ at 110 Hz). The plots for the flat and the temp condition are very similar, indicating that this subject cannot use the source of information provided by the temporal envelope fluctuations.[6] Best performance is obtained at $F0 = 172$ Hz, where thresholds decrease to values around 1.5%. Thresholds for the classic filter bank are in general very high, as was predicted from the plots in Fig. 3. An explanation for the good performance at /i/ 110 Hz could not be found. An artifact was not found, but cannot be excluded, given the complex nature of the stimuli.

Also, subject WM (Fig. 6) obtained in general noticeably lower thresholds for the triangle filter bank. The one exception at /i/ 145 Hz temp is probably due to the addition of a temporal cue. The results for subject JH (Fig. 7) show a different pattern for the flat and the temp conditions. Without a temporal cue, there is a clear advantage using the triangle filter bank. When the temporal cue is included, the thresholds for both filter banks are much closer, or even overlapping. The effect is very large for the classic filter bank, where thresholds drop considerably at almost all frequencies. The effect is much smaller for the triangle filter bank, which is not very surprising, given the already low thresholds obtained in the flat condition. A similar pattern is observed for subject SV (Fig. 8), where thresholds in the temp condition are lying closer together. The odd results at /i/ 110 Hz flat and /a/ 263 Hz flat could not be explained. In general, it

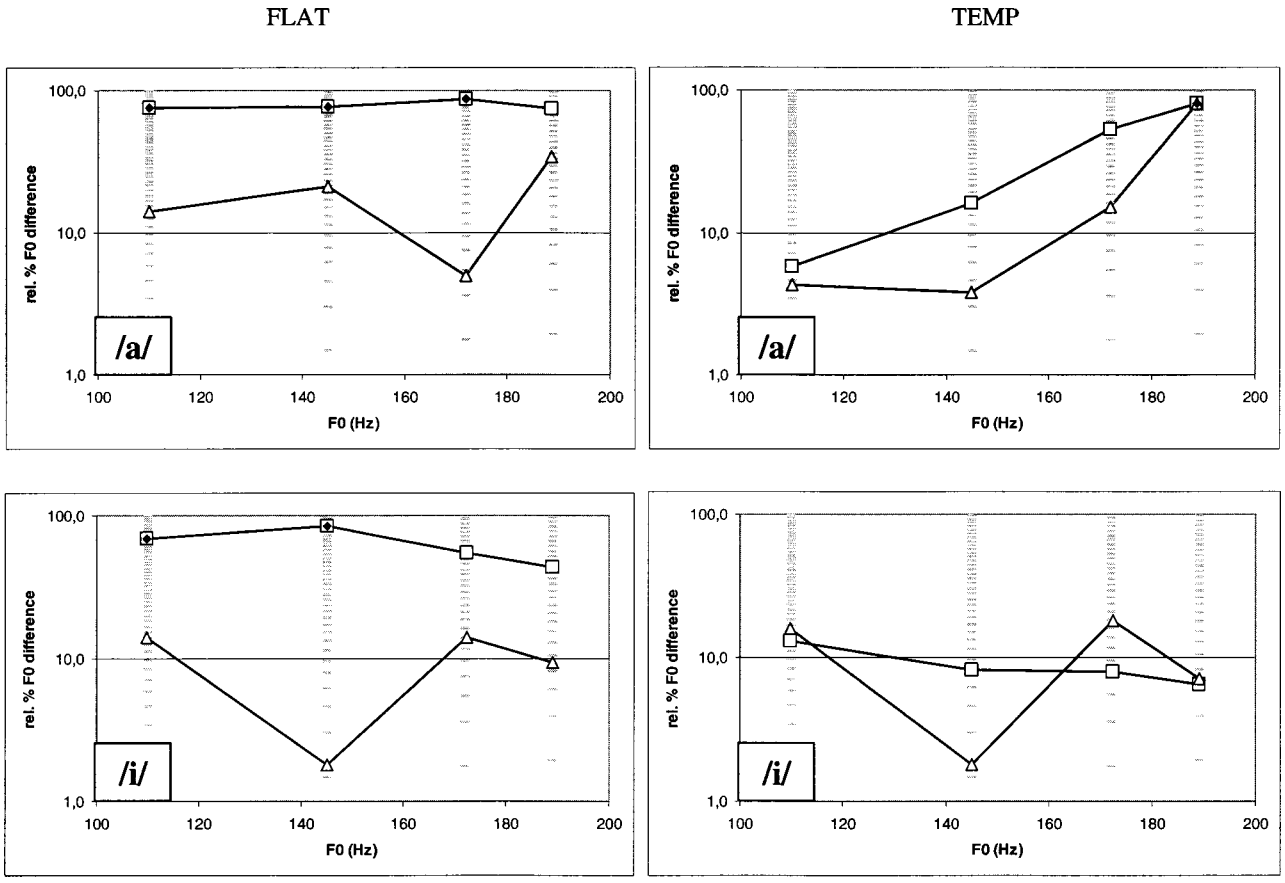


FIG. 7. Same as Fig. 5 for subject JH.

seems that a large improvement by inclusion of the temporal cue is obtained only if thresholds are very poor in the flat condition.

The data were analyzed using a four-factorial repeated-measures analysis of variance (ANOVA). The independent variables were filter bank (classic or triangle), presence of temporal fluctuations (flat or temp), vowel (/a/ or /i/), and standard $F0$ frequency (low or high). Since subject WM completed the tests at only two frequencies, the data at other frequencies for the other subjects were not included in the analysis. The main effect of the filter bank was significant ($p=0.002$). The main effects for the other factors were not significant: presence of temporal fluctuation ($p=0.135$), vowel ($p=0.398$), and standard $F0$ frequency ($p=0.348$). Only one interaction term was significant: vowel×standard $F0$ frequency ($p=0.031$).

To summarize, from the total of 56 measurements for each filter bank, the triangle filter bank outperformed the classic filter bank in 41 cases (flat: 25/28, temp: 16/28). If the results were separated less than two steps in the adaptive tracking, performance was judged to be equal, which occurred 11 times (flat: 1/28, temp: 10/28). The results for the classic filter bank were better in four cases (flat: 2/28, temp: 2/28). Also note that thresholds were below 2.5% in 15 cases of the triangle filter bank (never for the classic filter bank).

IV. DISCUSSION

There are two main differences between the two filter banks—the filter shape and the cutoff frequencies—and from the experiments it cannot be concluded which of those factors contributed the most to the better performance for the triangle filter bank. It is obvious that the classic filter bank is disadvantaged in the experiments, given the wide first filter. Note however that this is a common setting for clinical devices (Faulkner *et al.*, 2000). In general, the first filter is broad and does not always capture the fundamental component of a harmonic sound, due to the relatively high low-cutoff frequency, while the experiments illustrate the feasibility of using ''place-coding'' for the fundamental frequency. The triangle filter bank uses more electrode channels for the lowest frequencies, thereby ''smearing out'' the place coding of the fundamental component over a larger range in the cochlea: the same $\Delta F0$ will result in a larger difference in place of maximal excitation in the cochlea.

Regarding the filter shape, care should be taken when using a ''rectangular'' filter instead of a ''triangular'' filter. Analysis for the vowel /a/ as shown in Fig. 3 revealed that using the cutoff frequencies of the triangle filter bank, combined with a rectangular Butterworth shape, results in an irregular pattern, as shown in Fig. 9. The effect is due to the complex interplay between harmonics, resonant frequencies used for vowel synthesis, the root-mean-square leveling of the vowels, and the bandpass filters of the CIS strategy. The nonmonotonic behavior in the passband suggests suboptimal performance in a pitch-discrimination task, as is illustrated by the results of the experiments. This justifies the choice of a triangular shape.

A potential problem is that two different percepts are

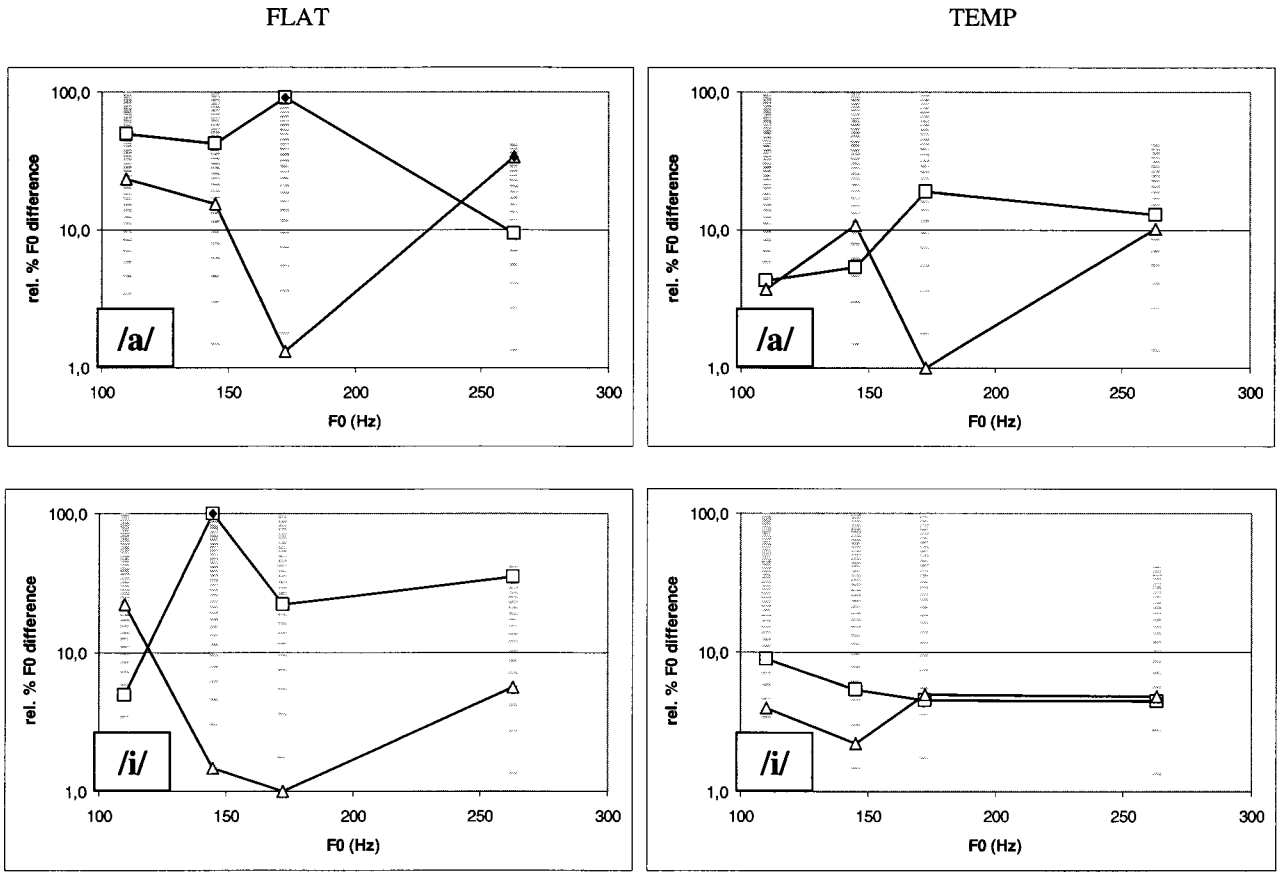


FIG. 8. Same as Fig. 5 for subject SV.

used to provide information on $F0$: rate and place pitch, which can be varied independently in a cochlear implant (McKay *et al.*, 2000). Yet, in normal hearing both mechanisms play a role: if $F0$ rises, the excitation pattern of resolved harmonics shifts basally and the nerve fibers are phase-locked to higher frequencies. Which of these mechanisms is the most dominant is still the subject of an ongoing debate (Moore, 1997). In currently available cochlear implant systems, place-coding is solely used to provide information on spectral shape. If place-coding is also used to encode $F0$, confusion may arise when both $F0$ and the spectral shape are changing, certainly if there is a lack of extensive training and/or experience with the new coding. Again, this also happens in normal hearing, and a cochlear implant system should mimic both mechanisms to narrow the gap between normal acoustic hearing and electric hearing.

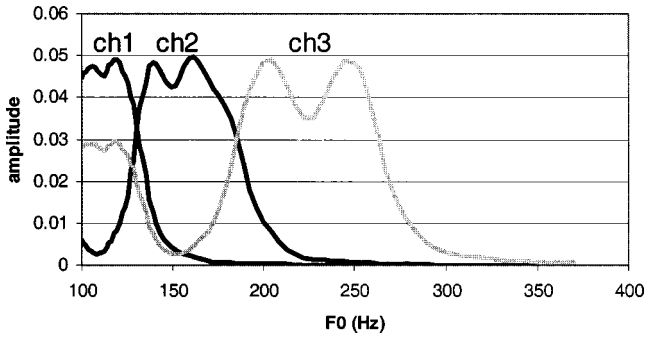


FIG. 9. The average amplitude in the first three channels (out of eight) as a function of the fundamental frequency for the vowel /a/, similar as in Fig. 3. The filters have a rectangular shape (Butterworth) as in the classic filter bank, but the cutoff frequencies are those of the triangle filter bank.

The proposed method should also work for very high $F0$'s, in the range where temporal mechanisms fail to provide information on $F0$. This temporal limit lies on the average around 300 Hz. Unfortunately, this could not be tested with the participating subjects, given their limited number of channels. At least 12 channels are needed to cover the speech frequency range with triangular filters (see Fig. 2). However, since all triangular filters are identical when plotted on a logarithmic scale, the observed output patterns as in Fig. 3 will simply shift to more basal electrodes for higher $F0$'s. So, as long as electrode discrimination ability does not change noticeably, subjects should have the same detection limits at high $F0$'s when using the triangle filter bank.

The triangle filter bank presented here only resolves the first harmonic of a complex sound, since all filters are one octave wide. From experiments with normal hearing, it seems that higher harmonics are more dominant for pitch perception (Moore, 1997; Dai, 2000), so smaller detection limits might be obtained if the triangle filter bank was extended to resolve higher harmonics. However, if more harmonics should be resolved, more filters—and thus more electrodes—are needed, and these filters should be smaller. The necessary filter width can be calculated using the following formula, in which NRH corresponds to the number of

resolved harmonics, and f_{BL} and f_{BH} to the lower and higher boundary frequencies, respectively:

$$\text{relative filter width} = \frac{f_{BH}}{f_{BL}} = 2^{1/\text{NRH}}. \quad (8)$$

As a consequence, the required number of filters increases linearly with the desired number of resolved harmonics (12 filters to resolve the first harmonic, 24 filters to resolve the second as well, ...). The experiments show, however, that $F0$ differences can already be detected by resolving the first harmonic only.

V. CONCLUSIONS

The experiments illustrate the feasibility of using ''place-coding'' to provide information on the fundamental frequency of a complex sound, including speech sounds. This place-coding is achieved by resolving the first harmonic in two adjacent filters, while the balance between both filter outputs is directly related to the frequency of the harmonic. The results show that detection thresholds for $F0$ differences in general decrease noticeably for the new filter bank, compared to a commonly used filter bank, and the effect is larger if no temporal cues are present in the stimuli. Differences between the results for both filter banks become smaller if the envelopes of the stimuli fluctuate according to $F0$, but a clear advantage is still observed for the new filter bank. In summary, the concept of the new filter bank can be used in current cochlear implant systems, thereby providing information on $F0$, supplementary to the information provided by the temporal envelope fluctuations. $F0$ difference detection limits will decrease, certainly at high frequencies where temporal cues will fall short.

ACKNOWLEDGMENTS

We would like to thank all subjects for their participation and Dr. Astrid van Wieringen for running part of the experiments. We also want to show gratitude to Koen Eneman from the ESAT laboratory of the KULeuven for his help in the filter design, Johan Laneau for his help in the statistical analysis, and Andrew Faulkner and an anonymous reviewer for their comments on the previous version of the manuscript. This study was supported by the Fund for Scientific Research—Flanders (Belgium), and partly by an OT project of the Research Fund of the KULeuven.

[1]MATLAB is a registered trademark of The Mathworks, Inc. (http://www.mathworks.com)

[2]LAURA was a registered trademark of Philips Hearing Implants, now Cochlear.

[3]COMPUTERISED SPEECH RESEARCH ENVIRONMENT, Version 4.5, AVAAZ Innovations, Inc. (http://www.avaaz.com)

[4]The procedure was altered for subject DD. From a prior experiment it was revealed that electrode 4, used for CIS channel 3 in the experiments, elicited a lower pitch than the deeper inserted electrode 3, used for CIS channel 2 (tonotopical order). This could result in a ''cue reversal:'' a stimulus with a low $F0$ can sound higher than a stimulus with a high $F0$. Therefore, the subject was initially instructed to indicate the stimulus with the highest pitch, as for the other subjects. However, if the adaptive procedure was diverging consistently ($\Delta F0$ grows), the experiment was restarted with the instruction to indicate the stimulus with the lowest pitch. This always yielded a converging adaptation path. There was no conflict with the temporal cue for this subject, since the results reveal that he did not seem to use the temporal cue (see further, Fig. 5).

[5]Subject WM only completed 16 conditions, due to an implant failure before all experiments could be carried out.

[6]This was confirmed in a prior experiment, where he performed at chance level when differentiating two sinusoidally amplitude-modulated pulse trains on one electrode channel with modulation frequencies of 150 and 250 Hz. The other subjects already had experience with pitch-discrimination tasks (Geurts and Wouters, 2001).

Dai, H. (**2000**). ''On the relative influence of individual harmonics on pitch judgment,'' J. Acoust. Soc. Am. **107**, 953–959.

Eddington, D. K., Dobelle, W. H., Brackman, D. E., Mladejovsky, M. G., and Parkin, J. L. (**1978**). ''Auditory prostheses research with multiple channel intracochlear stimulation in man,'' Ann. Otol. Rhinol. Laryngol. Suppl. **87**, Suppl. 53, 1–59.

Faulkner, A., Rosen, S., and Smith, C. (**2000**). ''Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: Implications for cochlear implants,'' J. Acoust. Soc. Am. **108**, 1877–1887.

Friedlander, B., and Porat, B. (**1984**). ''The modified Yule–Walker method of ARMA spectral estimation,'' IEEE Trans. Aerosp. Electron. Syst. **AES-20**, 158–173.

Geurts, L., and Wouters, J. (**2000**). ''A concept for a research tool for experiments with cochlear implant users,'' J. Acoust. Soc. Am. **108**, 2949–2956.

Geurts, L., and Wouters, J. (**2001**). ''Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants,'' J. Acoust. Soc. Am. **109**, 713–726.

Govaerts, G. (**1978**). ''Perceptuele structuren van synthetische en natuurlijke klinkers,'' Psychologica Belgica **18-1**, 27–67.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

McDermott, H. J., and McKay, C. M. (**1994**). ''Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea,'' J. Acoust. Soc. Am. **96**, 155–162.

McKay, C. M., McDermott, H. J., and Clark, G. M. (**1994**). ''Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees,'' J. Acoust. Soc. Am. **96**, 2664–2673.

McKay, C. M., McDermott, H. J., and Carlyon, R. P. (**2000**). ''Place and temporal cues in pitch perception: Are they truly independent?'' ARLO **1**, 25–30.

Moore, B. C. J. (**1997**). *An Introduction to the Psychology of Hearing*, 4th ed. (Academic, New York), edition.

Pijl, S., and Schwarz, D. W. F. (**1995**). ''Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes,'' J. Acoust. Soc. Am. **98**, 886–895.

Shannon, R. V. (**1983**). ''Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics,'' Hear. Res. **11**, 157–189.

Stevens, S. S. (**1957**). ''On the psychophysical law,'' Psychol. Rev. **64**, 153–181.

Tong, Y. C., and Clark, G. M. (**1985**). ''Absolute identification of electric pulse rates and electrode positions by cochlear implant patients,'' J. Acoust. Soc. Am. **77**, 1881–1888.

Townshend, B., Cotter, N., Van Compernolle, D., and White, R. L. (**1987**). ''Pitch perception by cochlear implant subjects,'' J. Acoust. Soc. Am. **82**, 106–115.

van Wieringen, A., and Wouters, J. (**1999**). ''Gap detection in single- and multiple-channel stimuli by LAURA cochlear implantees,'' J. Acoust. Soc. Am. **106**, 1925–1939.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (**1991**). ''Better speech recognition with cochlear implants,'' Nature (London) **352**, 236–238.

Construction and control of a physiological articulatory model

Jianwu Dang[a)]
School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan, and ATR Human Information Science Laboratories, Kyoto, Japan

Kiyoshi Honda
ATR Human Information Science Laboratories, Kyoto, Japan

(Received 16 May 2003; revised 17 October 2003; accepted 6 November 2003)

A physiological articulatory model has been constructed using a fast computation method, which replicates midsagittal regions of the speech organs to simulate articulatory movements during speech. This study aims to improve the accuracy of modeling by using the displacement-based finite-element method and to develop a new approach for controlling the model. A ''semicontinuum'' tongue tissue model was realized by a discrete truss structure with continuum viscoelastic cylinders. Contractile effects of the muscles were systemically examined based on model simulations. The results indicated that each muscle drives the tongue toward an equilibrium position (EP) corresponding to the magnitude of the activation forces. The EPs shifted monotonically as the activation force increased. The monotonic shift revealed a unique and invariant mapping, referred to as an *EP map*, between a spatial position of the articulators and the muscle forces. This study proposes a control method for the articulatory model based on the EP maps, in which co-contractions of agonist and antagonist muscles are taken into account. By utilizing the co-contraction, the tongue tip and tongue dorsum can be controlled to reach their targets independently. Model simulation showed that the co-contraction of agonist and antagonist muscles could increase the stability of a system in dynamic control.
[DOI: 10.1121/1.1639325]

PACS numbers: 43.70.Bk, 43.70.Aj, 43.70.Jt [AL] Pages: 853–870

I. INTRODUCTION

The production of speech involves the fine coordination of a serially ordered stream of changing articulatory, laryngeal, and respiratory movements. These skilled articulatory movements are concerned with well-practiced and overlearned muscular behaviors. A number of speech researchers have endeavored to discover the relationship between tongue movement and muscle activation via experimental approaches such as EMG experiments (Baer *et al.*, 1988; Dang and Honda, 1997). For the muscles involved in speech organs, however, experimental observations have succeeded for only a few large muscles, such as the extrinsic tongue muscles. The functions of the intrinsic muscles of the tongue have been investigated using tagged MRI (Niimi *et al.*, 1994), or tagged cine-MRI (Stone *et al.*, 2001), but the accuracy of such investigations is questionable, since a number of muscles are generally coactivated even to form a simple movement. Furthermore, it is often difficult to determine the mechanical load of a given muscle accurately because the load depends on the potential contribution of many other muscles. Therefore, a physiological model of the human speech organs is necessary for understanding the mechanism of speech production.

In the literature, a few physiologically based articulatory models have been reported. Perkell (1974) constructed the first computational model of the human tongue. His model was a two-dimensional (2D) projection of the tongue in the sagittal plane, which was composed of a lumped parameter and a lumped force system, equivalent to the finite-element method (FEM). The FEM approach has been applied to three-dimensional (3D) tongue models by Kiritani *et al.* (1976), Kakita *et al.* (1985), and Hashimoto and Suga (1986). These models were essentially based on infinitesimal elasticity methods, which describe the deformation process of the soft tissue as a sequence of quasistatic equilibrium configurations. Wilhelms-Tricarico (1995) proposed a method for modeling the tongue's soft tissue by large-scale FEM. Based on the method, he built a three-dimensional model of the tongue tissue and discussed the effects of geometric nonlinearities.

Hirai *et al.* (1995) developed a 2D physiological model unifying the tongue, jaw, and laryngeal structures (see also Honda *et al.*, 1994). The soft tissue of the tongue was formed by 2D FEM based on magnetic resonance imaging (MRI) data from a male speaker. The rigid organs such as the jaw, hyoid bone, thyroid cartilage, and cricoid cartilage were connected by muscles with elasticity. The dynamic balance of forces and moments was used as a mechanical principle to interface the soft and rigid structures. Because these structures were modeled independently, the model was computationally slow in achieving an equilibrium between the soft tissue and rigid bodies. Payan and Perrier (1997) reported a 2D biomechanical tongue model built by FEM. Their model produced V–V sequences according to the equilibrium point hypothesis (EPH), one of the common motor control theories (Feldman, 1986). The model demonstrated plausible movements for the tongue without incorporating jaw movement.

a)Electronic email: jdang@jaist.ac.jp

Sanguineti *et al.* (1998) employed a 2D model of the tongue, jaw, hyoid bone, and larynx to develop a control strategy based on the EPH (λ model). In their study, the dynamic behavior of the entire system was specified by its global kinetic and potential energy functions. They noted that the dynamic effects that occurred at the interface between the soft tissue and rigid organs were not negligible in modeling speech-like movements. This often resulted in serious stability errors because the dynamic processes of these two components were quite different.

In previous works (Dang and Honda, 2001; 2002), the authors measured the structure of speech organs based on volumetric MRI data obtained from a male speaker and constructed a partially 3D physiological articulatory model that consisted of the tongue, jaw, hyoid bone, and the vocal-tract wall. To obtain a time-efficient control for the physiological articulatory model, both the soft tissue and rigid organs were modeled as a mass–spring network that can efficiently and reliably simulate a large, fast deformation. A target-based control strategy was developed to generate muscle activation signals and realize the model's dynamic articulatory movements. The model was used as a synthesizer to generate speech sounds of short phrases (Dang and Honda, 1998; Dang *et al.*, 1999). It was later also used to estimate vocal-tract shapes from speech sounds (Dang and Honda, 2002). The model's disadvantage was that one of the important physical parameters for an elastic continuum, the Poisson ratio, was not taken into account, since the framework of the model consisted of mass points and volumeless springs. The revised model described in this paper has a marked improvement: the use of a viscoelastic cylinder to replace the volumeless spring. This enhancement accounts for the Poisson ratio and thus achieves a ''semicontinuum'' tissue model for the tongue body. In the present study, the new model is used to examine muscle functions for the tongue based on the model simulations. Furthermore, a control method is proposed to handle co-contraction during speech production.

II. MODELING USING DISPLACEMENT-BASED FINITE-ELEMENT METHOD

During natural speech, the tongue forms lateral airways by narrowing the tongue blade, or it makes a midsagittal conduit by grooving with bilateral tongue–palate contact, as seen in some consonants (e.g., /s/) and some vowels (e.g., /i/). A realistic model of the tongue must be capable of forming such midsagittal conduits and side airways, which are essential behaviors of the tongue in speech production. As a trade-off between computational cost and model verisimilitude, we have constructed a partial 3D model with a thick sagittal layer instead of a full 3D model.

A. Configuration of physiological articulatory model

The essential configuration of the model is the same as that used in the previous studies (Dang and Honda, 2001; 2002). The tongue model is a partial sagittal representation of a volumetric MR image of the tongue, which was obtained from a male Japanese speaker. According to our observations (Dang *et al.*, 1997), the tongue contacts the hard palate in the lateral area 1.5 cm from the midsagittal plane during most of the phonations. Such a contact interferes with the representation of the inherent characteristics of the tongue biomechanics. To achieve accurate modeling, the layer on each side of the midsagittal plane should be thinner than 1.5 cm in a model with two symmetric layers such as the proposed model of this study. Therefore, the lateral bound of the layers was set to be 1.0 cm apart from the midsagittal plane. The outlines of the tongue body are extracted from two sagittal slices: one is the midsagittal plane and the other is a plane 1.0 cm apart from the midsagittal on the left side. Under the assumption that the left and right sides of the tongue are symmetrical, the outline of the left side is copied to the right. The initial shape of the model adopts the tongue shape of a Japanese vowel [e], which approximates a centralized vowel in Japanese. Mesh segmentation of the tongue tissue roughly replicates the fiber orientation of the genioglossus muscle, the largest muscle in the tongue. The outline of the tongue body in each plane is divided into ten radial sections that fan out from the attachment of the genioglossus on the jaw to the tongue surface. In the perpendicular direction, the tongue tissue is divided concentrically into six sections. A 3D mesh model is constructed by connecting the section nodes in the midsagittal plane to the corresponding nodes in the left and right planes, where each mesh is a ''brick'' with eight corner nodes. Thus, the model represents the principal region of the tongue as a 2-cm-thick layer bounded by three sagittal planes. Figure 1 shows the initial shape of the tongue model based on the segmentation, with the surrounding organs. In this segmentation, the tongue tissue is represented as 120 eight-cornered brick meshes.

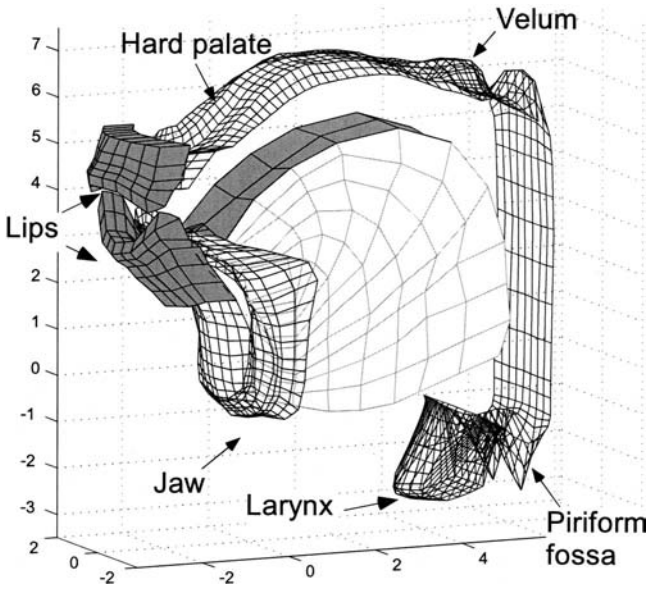


FIG. 1. Configuration of the physiological articulatory model.

To generate a vocal-tract shape, the articulatory model must include the tongue, lips, teeth, hard palate, soft palate (the velum), pharyngeal wall, and larynx. At the present stage, the lips and the velum are not modeled physiologically. They are included in the construction of vocal-tract shapes for speech synthesis but not in the generation of articulatory movements. The lips are defined by a short tube with a length and cross-sectional area, and the movement of

the velum is described by the opening area of the nasopharyngeal port. Outlines of the vocal-tract wall and the mandibular symphysis were extracted from MR images in the midsagittal and parasagittal planes (0.7 and 1.4 cm from the midsagittal plane on the left side). Again, assuming that the left and right sides are symmetrical, 3D surface models of the vocal-tract wall and the mandibular symphysis were constructed using mesh outlines with 0.7-cm intervals in the left–right direction. Figure 1 shows the model configuration of the vocal tract.

B. Modeling the tongue tissue using a truss structure

The soft tissue of the tongue has been commonly modeled using the finite-element method (FEM) (Kakita *et al.*, 1985; Wilhelms-Tricarico, 1995). The present study adopts displacement-based FEM as the basis of the modeling effort, which is referred to as an extended finite element method (X-FEM). The principal advantages of X-FEM are that the finite-element framework (sparsity and symmetry of the stiffness matrix) is retained and that a single-field variational principle is used (Zienkiewicz and Taylor, 1989; Belytschko *et al.*, 2002). Based on the principle of this X-FEM (see the Appendix for details), we can obtain the motion equation of equilibrium (1) for governing the linear dynamic response of a finite-element system

$$M\ddot{X}+B\dot{X}+KX=F, \tag{1}$$

where X, $\dot{X}$, and $\ddot{X}$ are the displacement, velocity, and acceleration vectors of the finite element assemblage, respectively. M, B, and K are the mass, damping, and stiffness matrices, respectively, in which the description is based on the displacements of the nodal points. F is the vector of externally applied loads.

1. Displacement function

The essence of X-FEM is that it uses the node displacements to describe the 3D deformation of a continuum. Exploiting this advantage, we reconsider the brick mesh mentioned above and then reduce some of the complexities in modeling the tongue tissue based on the common assumption that tongue tissue is an isotropic material.

In the model configuration, the basic mesh unit is described as an eight-node brick. Focusing mainly on the displacement of those nodal points, the relation of the relative locations of the nodes becomes a key point for analysis. For an isotropic material, this relation between the nodes can be easily represented by using an elastic solid to connect the nodes in all directions within the eight-node mesh. Based on this consideration, we use Hookean elastic bodies (Fung, 1993), referred to hereafter as cylinders, to connect all of the adjacent nodal points of the eight nodes in 3D. In the hexahedral mesh, there are 12 edges, corresponding to 12 cylinders. On the surface, there are two cylinders connecting the diagonal nodes crosswise on each of the six planes. Inside the brick, there are four cylinders transversely connecting the diagonal vertices. Altogether, each eight-node mesh is constructed from 28 viscoelastic cylinders. If the relative movement of every node pair can be described correctly by the changes in the cylinders, the node displacement in 3D can be fully described. Thus, the basic unit for the X-FEM analysis degenerates from the brick mesh to a truss. The displacement function for each truss is simplified as

$$H=\frac{1}{l_0}\begin{bmatrix} l_0-x & 0 & 0 & x & 0 & 0 \\ 0 & l_0-x & 0 & 0 & x & 0 \\ 0 & 0 & l_0-x & 0 & 0 & x \end{bmatrix}\begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix}, \tag{2}$$

where x is the distance from node i to node j, and l_0 is the length of the cylinder. T is a transformation matrix from a local coordinate (x, y, z) to the global coordinate (X, Y, Z)

$$T=\begin{bmatrix} \cos(\mathrm{X,x}) & \cos(\mathrm{X,y}) & \cos(\mathrm{X,z}) \\ \cos(\mathrm{Y,x}) & \cos(\mathrm{Y,y}) & \cos(\mathrm{Y,z}) \\ \cos(\mathrm{Z,x}) & \cos(\mathrm{Z,y}) & \cos(\mathrm{Z,z}) \end{bmatrix}. \tag{3}$$

Since each cylinder is a *Hookean elastic solid*, its deformation obeys Hooke's law. For an isotropic material, the stress–strain matrix is a 6×6 matrix described by Young's modulus E and Poisson ratio ν. Supposing that the Hookean body is a uniform cylinder with a radius r, its cross-sectional area varies uniformly when a force is loaded in the axial direction (x direction) alone. Thus, the 3D relation between the cylinder's length and its perpendicular dimension is simplified to

$$r=r_0\left(1-\nu\frac{l-l_0}{l_0}\right), \tag{4}$$

where l_0 and l are the lengths of the cylinder before and after applying a force and r_0 and r are the cylinder's radii corresponding to these lengths. At the end, the stress–strain matrix degenerates into a constant. When an axial force is loaded, the length variation of the elastic cylinder depends on the Young's modulus, while the change in the thickness obeys the law shown in (4). The deformation of a cylinder is transformed into the global coordinate by (2). Displacements of the nodes in a brick mesh are thus described by the deformation of the cylinders in the truss structure. Note that in the tissue model, each node connects seven cylinders (at the corners of the model) to 26 cylinders (inside the model). When a force is applied to a node, it is decomposed into seven to 26 directions corresponding to the axial direction of the connecting cylinders. Using the force decomposition, any force transmission within the model is achieved by a set of axial forces of the concerned cylinders. Therefore, we would not expect the simplification to introduce any significant error.

The dissipation caused by velocity-dependent damping must be taken into account when simulating the dynamic responses of the speech organs. For this purpose, we treat the cylinder as a viscoelastic body. According to Fung (1993), there are three types of models for representing a viscoelastic material: the Voigt model, the Maxwell model, and the Kelvin model. All three are composed of combinations of linear springs and a dashpot. The Voigt model consists of a spring parallel to a dashpot, the Maxwell model consists of a spring cascaded to a dashpot, and the Kelvin model is a combination of the two models. The relation of force F, displacement x and velocity $\dot{x}$ is described in (5a) for the Voigt model and (5b) for the Maxwell model

$$F = kx + b\dot{x}, \tag{5a}$$

$$x = \dot{F}/k + F/b, \tag{5b}$$

where k and b denote the stiffness and viscous coefficients, respectively. The Voigt model is good for describing a solid body, while the Maxwell model is good for liquid material. Therefore, we have adopted the Voigt model in modeling the tongue tissue. Comparing (5) and (1), one can also see that the Voigt model is easy to incorporate into the motion equation. When a force is applied on the Voigt model, a deformation gradually builds up as the spring shares the load. After the force is released, the dashpot displacement relaxes exponentially, and the original length is restored from the deformation.

2. Volume of cylinders

As described above, a hexahedral mesh with eight nodes consists of 28 cylinders. The tongue tissue is assembled from 120 such hexahedral meshes. The volume of the cylinders is determined by the basic principle that the volume summation of the cylinders concerning a mesh must be equal to the volume of the mesh.

In the assembled meshes, a cylinder can be shared by several adjacent meshes. In this case, the volume of the cylinder is assumed to distribute equally over the shared meshes. Thus, the weight coefficient of a cylinder for the relevant meshes is equal to the reciprocal of the number of the meshes sharing the cylinder. Among the 28 cylinders, there are four cylinders connecting the diagonal nodes inside a brick mesh, which are concerned with this mesh alone. Therefore, their weight coefficients are 1.0 in calculating the volume. The weight coefficient for the surface cylinders is 0.5 because they are shared by two adjacent meshes. The weight coefficient is 0.25 for the edge cylinders that are shared by four adjacent meshes.

The length of each cylinder is calculated in terms of the nodal coordinates of the two ends. To calculate the cross-sectional area of the cylinders, the shape of a mesh is first imagined as a uniform cylinder, whose length is the summation of the weighted length of all cylinders of this mesh. Then, the cross-sectional area of the uniform cylinder is determined by the quotient of the mesh's volume to the equivalent length. As a result, the thickness of the inside cylinders equals the cross-sectional area. The thickness for the surface cylinders is the summation of half the cross-sectional areas of two sharing meshes. Similarly, the thickness of the edge cylinders is the summation of the weighted cross-sectional areas of four cylinder-sharing meshes.

In this model, the viscoelastic cylinder is the basic element. The cylinder is a continuum that fully obeys physical laws. To reduce the effects of the discreteness of the truss-structure meshes, the Poisson ratio is also taken into account in the 3D meshes via a volume constraint that is described in the following section. Thus, this model can be thought of as a semicontinuum. This is one of the primary improvements from the previous version, where the meshes consisted of mass points and volumeless springs (Dang and Honda, 2001; 2002). With this improvement, the mass matrix becomes a consistent matrix, instead of the lumped mass matrix used in the previous version.

TABLE I. Parameters used in the present model.

Tongue Tissue	Density 1.0 g/cm^3	Young's modulus 20 kPa	Viscosity 2.0 kPa·s	Poisson ratio 0.49
Mandible	Weight 150 g	Young's modulus 9.6×10^6 kPa	Viscosity 9.6×10^5 kPa·s	⋯ ⋯
Hyoid Bone	Weight 5 g	Young's modulus 9.6×10^6 kPa	Viscosity 9.6×10^5 kPa·s	⋯ ⋯

3. Parameters and testing

In this study, the cylinder element is treated as a viscoelastic body. To describe the elastic properties, two essential parameters, Young's modulus E and Poisson ratio ν, are employed. A damping property parameter b is also introduced to describe the viscous property.

The Young's modulus influences the deformation of the body in the direction of the force, which basically measures the stiffness of the material. However, such mechanical parameters reported in past studies, which were obtained from different parts of the human body or from animals, have differed widely. Among them, Oka (1974) reported a value of 300 kPa for a contracted muscle. Duck (1990) found a value of 6.2 kPa for a human muscle under the rest condition, and a value of 110 kPa for the same muscle when contracted. Min *et al.* (1994) reported a Young's modulus of 20 kPa for the soft tissue of the vocal folds. Payan and Perrier (1997) used 15 kPa for the tongue tissue in their modeling study, and values ranging from 15 to 250 kPa for the tongue muscles corresponding to different levels of contraction. In this study, a value of 20 kPa is used for the Young's modulus in modeling the tongue tissue, following Min *et al.* (1994). For the viscosity, the viscous coefficient was 2 kPa·s, which was determined by a numerical experiment using this model. This value is about one-tenth that of the Young's modulus. Poisson ratio ν was set to 0.49, which is similar to that used in previous studies (Wilhelms-Tricarico, 1995; Payan and Perrier, 1997). The soft tissue of the tongue is considered to possess the same density as water. Therefore, a value of 1.0 g/cm^3 was used for the density of tongue tissue. These parameters are listed in Table I.

To verify the behaviors of the semicontinuum model with the Poisson ratio, we applied a force on a cuboid built on the truss structure and compared the Poisson ratio estimated from deformation of the cuboid with the original one. It was found that the estimated Poisson ratio varies with applied forces. To reduce this artifact, one more parameter, the ratio of the stress to the longitudinal strain, was employed. Duck (1990) showed that the Young's modulus depends to some extent on the ratio of the stress to the elongation, and is nearly constant for the soft tissue of the muscle if the relative elongation is less than 20%. The Young's modulus exponentially increases when the elongation ratio is larger than 20%. Following Duck, the Young's modulus is treated as a constant of 20 kPa in this study when the elon-

gation ratio is less than 20%, and its value increases as the elongation ratio becomes larger. The increasing rate of the Young's modulus functions as a parameter to maintain a constant Poisson ratio during tissue deformation.

The elongation-dependent Young's modulus was determined using a viscoelastic cuboid with a size of $6\times3\times2$ cm^3. The cuboid was divided into $5\times3\times2$ hexahedral meshes and constructed using the truss structure. When applying forces on the cuboid in the longitudinal direction, the Poisson ratio was evaluated by the changes of the cuboid in the axial and perpendicular directions using formula (4). An elongation-dependent Young's modulus was chosen to maintain the Poisson ratio around 0.49 while treating the cuboid either as a brick-assembled body or as a single brick. A numerical simulation was also conducted on the same cuboid by using ANSYS™ software, and the same behaviors were confirmed. The simulation showed that the truss-structure model based on X-FEM is sufficient for modeling the tongue tissue.

Compared with the traditional FEM, there are two major benefits of using this modeling. First, a fast computation is achieved without significant loss in accuracy. Second, the proposed model demonstrates excellent stability; no divergence was seen, even when a quite large force was loaded or when extreme deformation took place.

4. Volume constraint for tongue body

The tongue body is commonly considered to consist of incompressible tissue. However, the cubic meshes lack incompressible properties in the above model. If no volume constraint is taken into account, changes in the volume of the tongue were about 5% during tongue deformation. To reduce the changes, it is necessary to incorporate a constraint to maintain the volume of the tongue tissue. For this purpose, the Lagrange function first comes to mind. However, the constraint of the volume constancy introduced by the Lagrange function did not always work well; it sometimes interfered with tongue movement. One of the resulting phenomena, for example, was that on occasion the tongue could not move in response to a small change in force. It might be that the vectors for retaining constant volume are not distributed continuously over the multidimensional space consisting of the nodes' coordinates. For this reason, a procedure for minimizing volume changes is introduced to reproduce tissue incompressibility.

In the truss-structure model, the total volume of the tongue equals the summation of the volumes of all cylinders. Therefore, minimizing the changes in volume for all cylinders can achieve a volume constraint on the tongue body. When a force is applied on cylinder i in the axial direction, the change in the cylinder's volume is

$$\Delta V_i(X) = \pi r_0^2 l_0 - \pi r_0^2 \left(1-\nu\frac{\Delta l}{l_0}\right)^2 (l_0+\Delta l), \tag{6}$$

where the variation of the radius is represented by length increment Δl and Poisson ratio ν using (4). Using the Houbolt integration method (Bathe, 1996), motion equation (1) is rewritten in the finite difference expansions of $DX=B$, where D denotes the resultant matrix and B is the vector consisting of known terms in (1). The constraint is combined with the motion equation system by adding the volume difference and then minimizing the total error. Thus, the final system equations are derived from the following formula:

$$\frac{\partial}{\partial X}\left[\|DX-B\|^2+\alpha\sum_i \Delta V_i(X)^2\right]=0, \tag{7}$$

where α is the coefficient to adjust the tolerance of the volume changes in the tongue body. After introducing the volume constraint, the variation ratio of the volume of the tongue is reduced from about 5% to about 0.3%.

III. MUSCULAR STRUCTURE AND FORCE GENERATION

This model involves three kinds of external forces: muscle contraction, collision of the soft tissue with the rigid boundaries, and gravitational force. Muscular contraction is the source force to drive the model. Gravitational force always acts on all the nodes of the model in the vertical direction. The acceleration of gravity used in this model was 980 dyne·cm/s^2.

A. Modeling of muscles and rigid organs

The anatomical arrangement of the major tongue muscles was examined based on a set of high-resolution MR images obtained from the prototype speaker (Dang and Honda, 2001; 2002). The boundaries of the muscles were first traced in each slice of the MR images, and then superimposed on each other so that the contours for the major muscles could be identified. Thus, the genioglossus (GG) and geniohyoid (GH) were identified in the midsagittal plane, while the hyoglossus (HG) and styloglossus (SG) were mainly found in the parasagittal planes. The superior longitudinalis (SL) and inferior longitudinalis (IL) muscles were seen in both the midsagittal and parasagittal planes. The other intrinsic muscles such as the transversus and verticalis could not be identified in the MR images. The orientation of the tongue muscles was also examined with reference to the literature (Miyawaki, 1974; Warfel, 1993; Takemoto, 2001).

Figure 2 shows the arrangement of the tongue muscles used in the proposed model. Figure 2(a) shows the GG, which runs midsagittally in the central part of the tongue. Since the GG is a triangular muscle, and different parts of the muscle exert different effects on tongue deformation, it can be functionally separated into three segments: the anterior portion (GGa) indicated by the dashed lines, the middle portion (GGm) shown by the gray lines, and the posterior portion (GGp) indicated by the dark lines. The thickness of the lines represents the approximate size of the muscle fibers: the thicker the line, the larger the maximum force generated. Figures 2(b) and (c) show the arrangement of other extrinsic muscles, the HG and SG, in the parasagittal plane, where the thickest line represents the hyoid bone. In addition, two tongue-floor muscles, the geniohyoid and mylohyoid, are also shown in the parasagittal planes. The top points of the mylohyoid bundles are attached to the medial surface of the

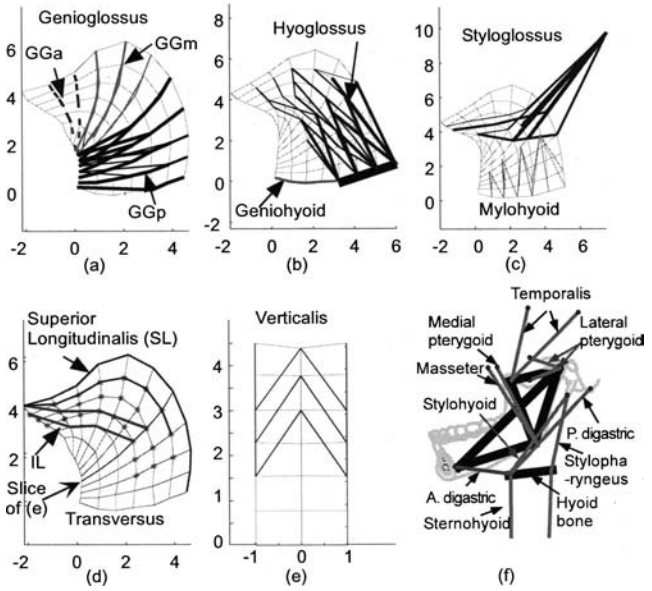


FIG. 2. Muscular structure of the model. (a)–(e) The tongue muscles in the midsagittal and/or parasagittal planes (dimensions in cm). (f) The mandible and hyoid bone complex.

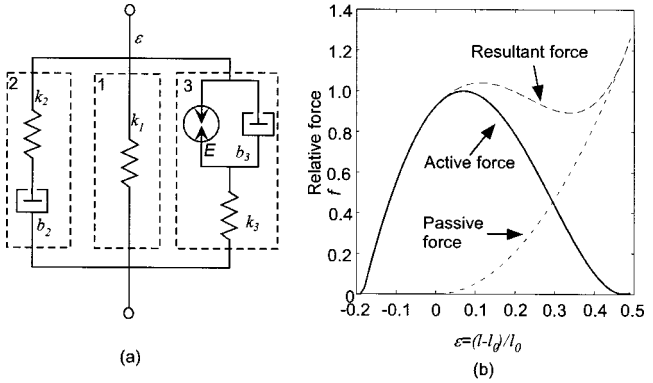


FIG. 3. Muscle modeling: (a) a general model of muscle unit: k and b are stiffness and dashpot, E is the contractile element; (b) generated force varies with stretch ratio ε.

mandibular body. All of the muscles are designed symmetrically on the left and right sides. Figure 2(d) shows the three intrinsic muscles of the SL, IL, and transversus. The transversus runs in the left–right direction, and its distribution is plotted in star markers. Figure 2(e) shows the structure of the verticalis muscle in a cross-sectional view sliced at the 5th section from the tongue floor, shown in Fig. 2(d). Altogether, 11 muscles are included in the tongue model.

Figure 2(f) shows the model of the jaw–hyoid bone complex. The right half of the mandible is drawn in the background with the pale gray lines. The model of the jaw has four nodes on each side, which are connected by five rigid beams (thick lines) to form two triangles with a shearing beam. These four points, which are similar to those used by Laboissière *et al.* (1996), are selected as the attachment points for the jaw muscles. The jaw model is combined with the tongue model at the mandibular symphysis. The hyoid bone is modeled as three segments corresponding to the body and bilateral greater horns. Each segment has two nodes connected with a rigid beam. Eight muscles indicated by thin lines are incorporated in the model of the jaw–hyoid bone complex, where the structure of the muscles is based on the anatomical literature (Warfel, 1993). The small circles indicate the fixed attachment points of the muscles. Since the rigid organs below the hyoid bone, such as the thyroid and cricoid cartilages, are not included in the present model, two viscoelastic springs are used as the strap muscles. The temporalis and lateral pterygoid are modeled as two units to represent their fan-like fiber orientation. The digastric muscle has two bellies, anterior and posterior, and is modeled to connect the hyoid bone at a fixed point. All of these muscles are modeled symmetrically left and right. Jaw movements in the sagittal plane involve a combination of rotation (change in orientation) and translation (change in position). Although there is no one-to-one mapping between muscle actions and kinematic degrees of freedom, the muscles involved in the jaw movements during speech can be roughly separated into two groups: the jaw-closer group and the jaw-opener group.

According to Duck (1990), the Young's modulus of the bones is $11.5-12.0\times10^6$ kPa for the human femur and 6.9×10^6 kPa for the human tibia. In this study, we used the value obtained from the human tibia in modeling the bony structures, the mandible and the hyoid bone. For the weight of the bones, Yamazaki (1933) investigated the weight of the cranium and mandible using 92 dry skulls from Japanese specimens. His result showed that the weight of the male jaws was around 90 g. According to this literature, the equivalent mass of the living jaw is roughly estimated to be 150 g including water and surrounding tissue. To evaluate the mass for the hyoid bone, the structure of the hyoid bone was extracted and measured using volumetric computer topographic data. The volume of the hyoid bone was about 2.5 cm^3 for a male subject. Based on this measurement, an equivalent mass was set at 5 g for the hyoid bone. The masses are uniformly distributed over the hard beams. To provide a uniform computational format, rigid beams were also treated as viscoelastic links with a high Young's modulus so that they could be integrated with the soft tissue in the motion equation. The viscosity was set to be about one-tenth of the Young's modulus in value, the same ratio as that used for the soft tissue. Note that the viscosity for the rigid beams is less important because the muscle force is not strong enough to cause any definite deformation on the rigid beams.

B. Generation of muscle forces

In formulating a generalized model of the muscle, this study adopts a commonly accepted assumption: a force depending on muscle length is the sum of the passive component (independent of muscle activation) and the active component (dependent on muscle activation). Figure 3(a) shows a diagram of the rheological model for a muscle sarcomere (Morecki, 1987), which is an extended model of Hill's model (Hill, 1938). The muscle model consists of three parts that describes the nonlinear property, the dynamic (force-velocity) property, and the force–length property. The properties of the muscle sarcomere can be described by a set of differential equations

$$\sigma_1 = k_1 \varepsilon,$$

$$\frac{\dot{\sigma}_2}{k_2} + \frac{\sigma_2}{b_2} = \dot{\varepsilon},$$

$$(\sigma_m + \sigma_3)(k_3 + E) + \frac{k_3 d(\sigma_m + \sigma_3)}{dt} = b_3 E \dot{\varepsilon} + E k_3 \varepsilon, \tag{8}$$

$$\sigma = \sigma_1 + \sigma_2 + \sigma_3,$$

where σ_1, σ_2, σ_3 are the stresses of each part, and σ is the total stress of the sarcomere; $\sigma_m = E\varepsilon^2$, and $\varepsilon = (l - l_0)/l_0$. l is the current length of the muscle sarcomere, and l_0 is the original length of the muscle sarcomere at rest position of the tongue.

The first three equations in (8) describe parts 1, 2, and 3 of the muscle model. Part 1 is a nonlinear spring k_1, which is involved in generating force only when the current length of the muscle sarcomere is longer than its original length. The value of k_1 is selected as $k_1 = 0.05 k_0 \varepsilon$, where $\varepsilon > 0$ and k_0 is the stiffness of the tongue tissue. Part 2 consists of a Maxwell body and is always involved in force generation. According to Eq. (5b), the force generated by this part is determined by two factors: the velocity of the muscle length and the previous force of this branch. As shown in the literature (cf. Zajac, 1989; Wilhelms-Tricarico, 1995; Laboissière *et al.*, 1996), the force–velocity characteristic of the muscle is treated as independent of the previous force. To emphasize the effect of the velocity of the muscle length, a relatively larger stiffness and a smaller viscous component are used in this part. The value of k_2 was set to be twice that of the tongue tissue, while b_2 was on the order of one-tenth that used in the tongue body.

Part 3 of the muscle sarcomere corresponds to the active component of the muscle force, which is the Hill's model consisting of a contractile element parallel to a dashpot and then cascaded with a spring. This part generates force as a muscle is activated; its characteristics are described by the third equation. In model computations, however, we use a force–length function of the muscle tissue instead of the third equation. The force–length function was derived by matching the simulation and empirical data using the least-square method (Morecki, 1987). The function arrived at a fourth-order polynomial of the stretch ratio of the muscles

$$\sigma_3 = 22.5\varepsilon^4 + 3.498\varepsilon^3 - 14.718\varepsilon^2 + 1.98\varepsilon + 0.858, \tag{9}$$

which has a similar shape to that used by Wilhelms-Tricarico (1995). This empirical formula is valid for $-0.185 < \varepsilon < 0.49$. The active force is assumed to be zero if ε is out of the given range. Figure 3(b) shows the relationship between the stretch ratio of the muscle sarcomere and the generated force including the passive force. This figure demonstrates the force–length characteristic of the muscle model.

Since a muscle consists of a number of muscular fibers with various lengths and thicknesses, the general lumped rheological parameters of the muscle tissue are not sufficient for determining the muscle-generated force. For this reason, we introduced a parameter, the ''thickness'' of the muscle fiber, into the force generation. The thickness works as a coefficient for all three parts of the muscle sarcomere, which ranged from 0.1 to 4. The value for a given muscle is determined by making the maximum force of the muscles consistent with empirical data (Laboissière *et al.*, 1996; Sanguineti *et al.*, 1997). As shown in Fig. 2, for example, the GGp is thicker than the GGa; so the GGp generates a stronger force than the GGa.

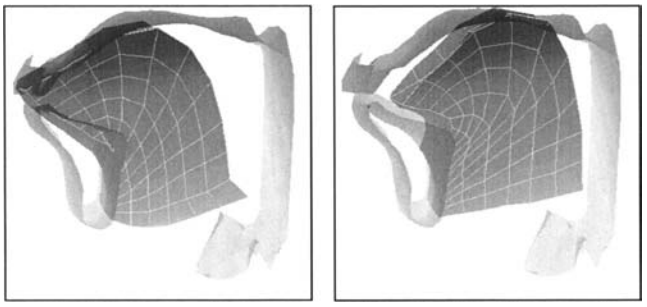


FIG. 4. Tongue deformation and movement from the initial position 20 ms after applying forces on GGp (left) and SG (right), where the jaw-opener muscle was activated in both cases. The force on the tongue muscles was 4 N, and 3 N for the jaw-opener muscle.

C. Evaluation of the articulatory model

Having completed the construction of the proposed articulatory model, including the soft tissue, rigid organs, and muscular structure, we conducted several numerical experiments to evaluate the characteristics of the model.

Figure 4 shows examples of the movements of the tongue and jaw driven by a force of 4 N for a tongue muscle and 3 N for the jaw-opener muscles, where the computation step was 5 ms in this simulation. The left panel demonstrates the tongue shape derived from the rest posture 20 ms after force was applied on the GGp and the jaw opener. The tongue tip and tongue dorsum moved upward and forward about 0.5 and 0.3 cm, respectively. The jaw lowered 0.16 cm. In most cases, tongue movement had a high positive correlation with jaw movement. An opposite movement was set for the tongue and jaw in order to test some of the extreme articulations that might occur in emphasized speech (Erickson, 2002). When the tongue and jaw moved in the opposite directions, the bilateral sides of the tongue blade contacted with the hard palate within 20 ms. The collision of the tongue and surrounding wall was not considered in this testing.

The right panel shows the posture of the model derived from the rest position 20 ms after applying a force of 4 N on the SG and 3 N on the jaw opener. Displacements were about 0.16 cm for the jaw and 0.8 and 0.6 cm for the tongue tip and tongue dorsum, respectively. The corresponding velocity for the control points was 41 cm/s for the tongue tip, 31 cm/s for the dorsum, and 8 cm/s for the jaw. For comparison, the authors measured the maximum velocity of the articulators from the articulographic data obtained from three speakers (Okadome and Honda, 2001). The maximum speeds were 40 cm/s for the tongue tip, 32 cm/s for the dorsum, and 12 cm/s for the jaw. The velocity was about the same for both the tongue tip and the dorsum in the measurement and model simulation, while the velocity of the jaw in this example was smaller than that of the measurement. A similar measurement

TABLE II. Comparison of the velocities (cm/s) of the tongue tip, tongue dorsum, and mandible obtained from simulation and observations.

	Tongue tip	Tongue dorsum	Mandible
Model	41	30 (18)[b]	8
EMMA	40	32 (22)[b]	12
Stevens (2000)[a]	32	15[b]	⋯

[a]Measured from the figures (Stevens, 2000).
[b]Vertical movement alone.

in the literature (Stevens, 2000) is cited in Table II and was used to evaluate the model simulation. The comparison showed that our simulation of the velocities was consistent with those results for the tongue tip and the tongue dorsum.

In Fig. 4, one can see that the meshes in the tongue demonstrated well-balanced deformation. In the right panel, the tongue dorsum made a closure with the palate within 20 ms, and the closure was maintained well until the jaw reached its stationary position with an approximately 1.0-cm aperture. The same test was also conducted in some extreme cases of applying a large force of 7 N on a number of tongue muscles. For the given forces, the tongue had some extreme deformations, but the deformation was smooth and no divergence was seen. In these evaluations, the proposed model demonstrated excellent performance on both static and dynamic behaviors.

D. Collision of the tongue and the vocal-tract wall

In speech articulation, the tongue often contacts the teeth, hard palate, and jaw as it moves. The tongue tip, for example, collides with the vocal-tract wall when producing consonants such as /t/ and /l/. The lateral parts of the tongue contact the hard palate to form a narrow airway of the vocal tract in producing the vowel /i/ or alveolar fricatives. Thus, the contact of the tongue with the vocal-tract wall is one of the important factors in achieving accuracy and stability in the dynamic control of the tongue. It also generates external forces that affect tongue deformation when the contact occurs. Therefore, the realization of tongue–wall contact is an essential task for a physiological articulatory model.

Since the shape of the vocal-tract wall is too complex to be described by an analytic function, the contact of the tongue with the vocal-tract wall, unlike the other constraints, cannot be combined into the motion equations systematically. As an alternative, we propose a method with three steps to deal with the contact between the tongue and the tract wall to check whether or not the nodes of the tongue cross through the tract wall; find the equilibrium position on the wall for the nodes outside of the vocal tract, and distribute the collision force. If a node crosses through the wall of the vocal tract during articulation, its trajectory must have an intersection with the tract wall. Since the tract wall was assembled by triangle planes, we first identify the plane with which the trajectory intersected on the wall and then calculate the collision force of the node when it is bounded on the wall. The following formula is used to estimate the collision force:

$$f_x = \sum_i (k_i \Delta l_{xi} + b_i \Delta l_{xi}/h), \tag{10}$$

where i is the index of the cylinders connected to the node, and h is the computation step. k_i and b_i are the stiffness and viscous coefficients of cylinder i. Δl_{xi} is the increment of cylinder i in the x dimension caused by the wall bounding. f_x is the x component of the resultant bounded force. Using the same approach, the force can be calculated for the y-dimension f_y and z-dimension f_z. To reach an equilibrium position on the wall, the coordinates (x, y, and z) of the node must meet the following simultaneous equations:

$$\begin{aligned} &A(x-p_x)+B(y-p_y)+C(z-p_z)=0,\\ &\alpha(x-p_x)=f_x,\\ &\alpha(y-p_y)=f_y,\\ &\alpha(z-p_z)=f_z, \end{aligned} \tag{11}$$

where p_x, p_y, and p_z are the coordinates of the intersection on the plane; A, B, and C are the norm of the plane; and α is an unknown equivalent factor of the stiffness and viscosity. Since such a deformation cannot be predicted in the motion equations, the collision forces must be considered additionally. In model calculation, the collision force above is taken into account at the next computation step as an input.

When the tongue slides on the wall surface to reach an equilibrium position, friction between the tongue and the palate is a considerable factor. This friction can be approximated as the force generated in laminar flows of the mucosa, since there is much mucosa on the surfaces of the tongue and the hard palate. Such a friction force is proportional to the viscosity of mucosa, the contact area between the tongue and palate, and the velocity of the tongue, but is inversely proportional to the thickness of the mucosa. To obtain the order of the magnitude for such friction, we roughly estimate the friction force based on the following conditions. Suppose the mucosa has about the same viscous coefficient as water, which is 0.01 dyne·s/cm^2 at 20 °C, and the thickness of the mucosa between the tongue and palate is very thin, about 0.01 cm. The maximum velocity of the tongue is about 40 cm/s (see Table II). Under the given conditions, the force caused by the friction for a unit area is about 40 dynes, or 4×10^{-4} N. This friction force is much smaller than the forces caused by muscle contraction and wall reaction. Therefore, this friction is ignored at the current model. Since the friction can be expected to stabilize the tight constriction to some extent, it may be helpful in achieving accurate control.

IV. ESTIMATION OF MUSCLE FUNCTION BY MODEL SIMULATIONS

As stated in the Introduction, many studies have sought to ascertain the relationship between tongue movement and muscle activation. However, it is generally difficult to determine such a relationship from experimental observation alone because critical parameters such as mechanical load on a muscle cannot be observed. This section examines muscle functions using model simulations.

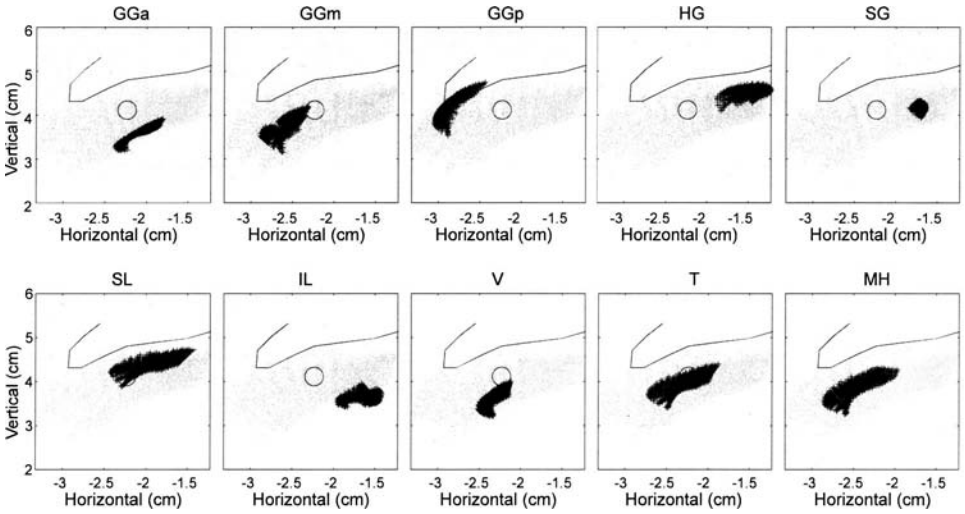


FIG. 5. Starting point region of the tongue tip, shown by the pale small circles, and the ending points driven by a force of 4 N with a 150-ms duration, shown by the dark crosses.

A. Muscle activations and tongue movements

In order to evaluate tongue muscle functions by simulation, tongue movements are represented using two points of the tongue tip and the tongue dorsum, which are referred to as the *control points* (Dang and Honda, 2001; 2002).

1. Function of tongue muscles

In producing an utterance, the tongue can start its movement from many different positions or shapes. The coming deformation of the tongue involves all of the past deformations and/or the history of the muscle forces to some extent. To account for the effects of the past deformations on tongue behaviors, a scattered region of starting points was designed for the control points, based on observations obtained from the prototype speaker of the model using the x-ray microbeam system (Hashi *et al.*, 1998). The control points were first moved from the initial position to a given starting point. The initialization movements were governed by four muscle combinations: GGp–GGm–GGa, GGp–SG–GGa, HG–GGm–GGa, and HG–SG–GGa. Activation forces for each muscle were set to be six levels between 0 and 6.0 newtons with a 120-ms duration. Nonlinear intervals were adopted in the six levels to achieve a uniform distribution. As a result, 864 starting points were generated for each control point, as shown by the small gray circles in Figs. 5 and 6 as a reference.

After the control points arrived at the given starting points, all forces were released and then the tongue body was driven by a specific muscle with a given force. The force was 4 N with a 150-ms duration. Figure 5 plots the ending locations of the tongue tip, shown by the dark crosses. The small gray circles show the starting points, and the large circle denotes the initial location of the control point. Note that the boundary and contact force of the tract wall were not considered in estimating the muscle functions.

As shown in this figure, even though the tongue tip started from widely scattered points, it converges to a specific, small region for all of the muscles. When the GGa contracts, the tongue tip concentrates to a strip of a region that is lower than the initial position, shown by the large circle. This means that the function of the GGa is to lower the tongue tip. Similarly, the GGm moves the tongue tip forward and downward, while the GGp drives the tongue tip forward and upward. The HG moves the tongue tip backward and slightly upward. The SG mainly drives the tongue tip backward, where the convergence region for the SG is smaller than that of the others. This implies that the SG has

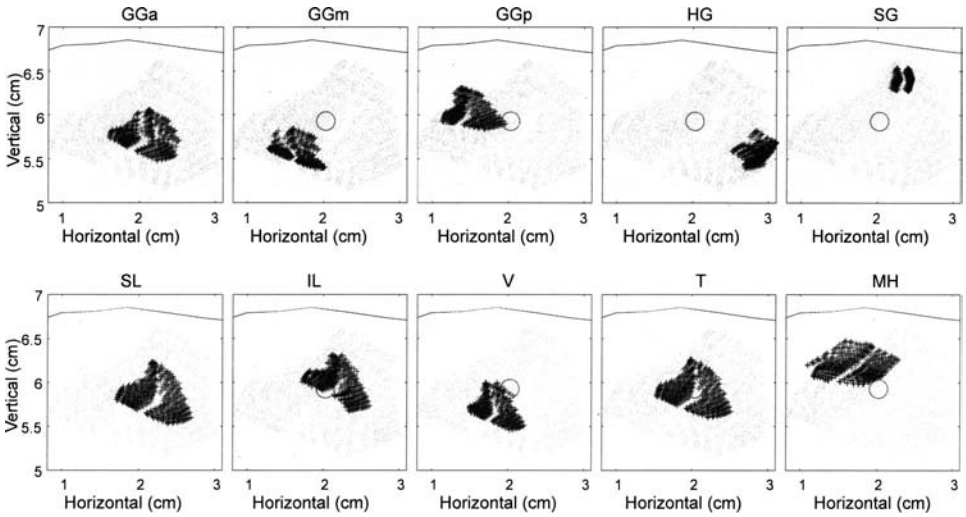


FIG. 6. Starting point region of the tongue dorsum, shown by the small gray circles, and the ending points driven by a force of 4 N with a 150-ms duration, shown by the dark crosses.

a clearer equilibrium position for the tongue tip, or that this equilibrium position is reached faster than the others.

For the intrinsic muscles, the inferior longitudinalis (IL) has a definite function that moves the tongue tip backward and downward, while the superior longitudinalis (SL) drives the tongue tip upward and backward. The verticalis (V) moves the tongue tip downward. The transversus (T) and mylohyoid (MH) did not show any definite contribution to movement of the tongue tip. The geniohyoid is not shown in the figure because its contribution to both the tongue tip and the dorsum was small, or underestimated in the model simulation.

Figure 6 shows the scattering of the starting points and ending locations for the tongue dorsum, where the simulation condition was the same as that for Fig. 5. Among the extrinsic muscles, the GGm, GGp, HG, and SG show definite functions that drive the dorsum to go front-lower, front-upper, back-lower, and back-upper, respectively. These results are consistent with the findings in previous studies (cf. Baer *et al.* 1988), while the GGa did not show any significant effect on the dorsum movement. The MH has a certain effect on the tongue body, which drives the tongue dorsum upward and forward. The verticalis moves the dorsum downward and forward to a certain extent, while the IL moves the dorsum backward and upward. The SL moves the dorsum slightly backward and downward, while the transversus did not show any contribution to the dorsum movement.

Putting Figs. 5 and 6 together, one can see that when the HG is activated the tongue dorsum moves backward and downward, while the tongue tip moves backward and slightly upward. This suggests that an accompanying rotation of the tongue body takes place during such movements. The above simulation shows that when the same muscle forces are applied on the tongue, the control point from different starting positions converges to a region but not a point during the given duration. Since the duration is close to vowel duration in speech with a normal rate, this suggests that the observed vowel target regions in the electromagnetic articulographic data (cf. Dang *et al.*, 2002) may be a consequence of tongue biomechanics. The simulation results also suggest that the tongue muscles may need a longer time to achieve their equilibrium position.

2. Muscle forces and equilibrium position

The results plotted in Figs. 5 and 6 show that the control points were driven to arrive at specific regions in spite of their widespread starting points. For the purpose of model control, we consider two issues. One is whether the control points equilibrate at one point or converge to a sufficiently small region when the activation duration is sufficiently prolonged. The other is whether the relationship between a muscle force and an equilibrium position is unique.

A numerical experiment was conducted to answer these two questions. Four starting points were generated by four muscle combinations with a force of 5.0 N for all muscles. The tongue was then driven from the starting points by a specific muscle with a 400-ms duration. The activation forces were 1.0, 2.0, 3.5, and 5.0 N. This simulation was carried out on all of the tongue muscles. Figure 7 shows

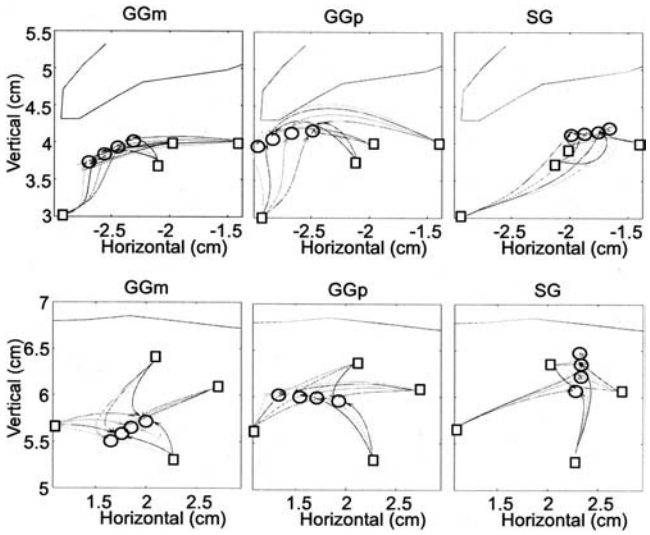


FIG. 7. Trajectories of the tongue tip (upper) and the dorsum (lower) from distinct starting points (indicated by squares) to certain ultimate locations (shown by circles) that correspond to the force levels of 1, 2, 3.5, and 5 N.

examples for three extrinsic muscles. The upper panels demonstrate the results for the tongue tip, and the lower panels are for the tongue dorsum. For each starting point, the trajectories of the control points spread out along different paths corresponding to the force levels. For different muscles, the trajectories have dissimilar curvatures. When activating the GGp, for example, both the tongue tip and the dorsum tended to reach force-dependent locations in a straight path. In the cases of the GGm on the dorsum and the SG on the tongue tip, the control points move along curved paths from the starting points to the final locations. The difference between the paths is mainly caused by previous deformation (starting point) and/or the history of the forces. However, they finally converge at one point, where the model reaches an equilibrium state. In the simulation, the control points generally reach their equilibrium positions within about 300 ms. These final points are referred to as the *equilibrium positions* hereafter. It is interesting to find that the equilibrium position of each muscle shifts monotonically as the force level increases. This means that the equilibrium position and the force level have a unique relation for a given muscle.

B. Mapping between equilibrium position and articulatory target

In order to control the model via the muscle forces, it is useful to find a mapping between the muscle forces and articulatory targets. Such a mapping is constructed in this section based on the equilibrium positions.

1. Equilibrium positions of muscles

As shown in Fig. 7, the equilibrium position (EP) for each muscle corresponds to its activation level, despite past deformations. This relation provides a connection between a muscle force and a spatial point in the articulatory space which is invariant for a given muscle structure. Using such a connection, a unique mapping can be obtained from a muscle force to a spatial position. However, the inverse mapping

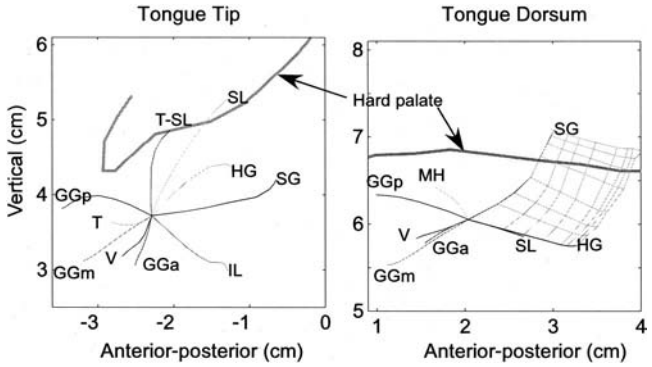


FIG. 8. Coordinates consisting of the equilibrium positions corresponding to the activation forces ranged between 0 and 6 N. The net in the right panel consists of the contour lines of the EPs of SG and HG.

TABLE III. Contribution factors and the major contributions of the tongue muscles.

Muscle names	Tongue tip	Tongue dorsum	Contribut. (>0.3)
Genioglossus A.	0.37	0.39	Both
Genioglossus M.	0.58	0.72	Both[a]
Genioglossus P.	0.68	0.74	Both[a]
Hyoglossus	0.74	0.86	Both[a]
Styloglossus	0.94	1	Both[a]
Superior longitudinalis (SL)	1	0.45	Tip[a]
Inferio longitudinalis	0.69	0.13	Tip[a]
Verticalis	0.36	0.43	Both
Transversus (T)	0.31	0.22	None
T-SL group	0.65	0.17	Tip[a]
Myohyoid	0.21	0.36	Dorsum

[a]Shows the muscles whose contribution factor is larger than 0.5.

from spatial position to muscle force is not unique, which brings about the ''one-to-many'' problem. Note that the *equilibrium position* depends on the muscle activation level alone, where no hypothesis is involved with the threshold of the stretch reflex. To distinguish the proposed method from the *equilibrium point hypothesis* (EPH, λ model), we use the term of *equilibrium position* hereafter, but not *equilibrium point*.

In order to obtain a series of EPs, each muscle is activated by an altering force for 300 ms, where the force was applied at levels of 0.0, 0.1, 0.2, 0.4, 1.0, 2.5, 4.0, and 6.0 N, respectively. Thus, a coordinate is established based on the EPs for each control point. Figure 8 shows the coordinates for the tongue tip and tongue dorsum. Since the EPs shift monotonically, the equilibrium position for a muscle can be expected to move along the path consisting of the EPs as the muscle force varies continuously, as long as the other muscles' forces remain unchanged. The path built on the EPs can be considered a ''vector'' that spreads out from the rest position. The HG moves the tongue tip backward and slightly upward as the activation level increases. The SG drives the tongue tip backward almost horizontally, while the GGa draws the tongue tip downward. The tongue tip is driven forward-upward by the GGp and forward-downward by the GGm. The intrinsic muscles SL and IL move the tongue tip largely upward-backward and downward-backward, respectively.

In Fig. 8, one can see that if the T-SL muscle group is not considered, there is a large blank space between the vectors of the SL and GGp in the tongue tip coordinate. This means that no single muscle can move the tongue tip in that direction. Since this region is important for constructing the alveolar consonants, we have proposed a muscle group consisting of the transversus (T) and the SL to fill this blank space with a new EP vector of T-SL. After this muscle group is added, the EP vectors are distributed almost uniformly in the coordinate for the tongue tip. Such a coordinate makes it possible to move the tongue tip involved in any direction.

The right panel of Fig. 8 shows the EP vectors for the dorsum. Unlike the tongue tip, the dorsal coordinate has a much larger scale in the horizontal direction than in the vertical direction. The extrinsic muscles (except the GGa) have definitely larger EP vectors than the others. The EP vectors of the GGp and HG show diametrically opposing directions. This indicates that the GGp and HG work antagonistically in governing the tongue dorsum. Similarly, the GGm and SG form another antagonist muscle pair. Therefore, there are obviously two large antagonist muscle pairs controlling the dorsum.

As the dorsum is driven upward and backward by the SG, it begins to cross through the hard palate when the muscle force reaches 1.0 N. The simulation shows that the model can make the gesture for /k/ with a force of about 1.0 N when the jaw is in the rest location. For the GGp, its EP vector did not reach the hard palate in the dorsal coordinate, while a contact between the tongue dorsum (or the blade) and the anterior portion of the hard palate was required in a number of articulations. As shown in Fig. 4, the tongue dorsum actually had reached the hard palate within 20 ms in the simulation, but the control point (the fifth node from the tongue tip) could not reach the palate. This is because the highest point of the dorsum did not always correspond to a fixed node of the model during the deformation, as demonstrated in the left and right panels of Fig. 4. This study uses the fifth node to represent the dorsum since the highest point is around the node in the most cases.

2. Construction of the EP map

In the proposed model, there are 11 muscles involved in the tongue body. To simplify descriptions, this study focuses on the major muscles only. For this purpose, muscle contributions are evaluated using the amplitude of the EP vectors, that is, the shift distance of the EP as the force increases from 0 to 6 N.

Among the EP vectors, the SL has the largest amplitude in the tongue tip coordinate, while the SG generates the largest vector for the tongue dorsum, as shown in Fig. 8. The vector amplitudes are normalized by the maximum one for the tongue tip and dorsum, respectively. The normalized amplitudes reflect a relative contribution of the muscles to tongue movements, and they are referred to as the *contribution factor*. Table III shows the contribution factor of the muscles. A muscle is considered to have a certain contribution to the tongue movement if its contribution factor is larger than 0.3. The superscripted [a] denotes the muscles with a contribution factor larger than 0.5. As a result, the most

extrinsic muscles (except the GGa) have larger power over both the tongue tip and tongue dorsum. Among the intrinsic muscles, the SL and IL show definite contributions to the tongue tip. The transversus demonstrates a certain contribution when it is grouped with the SL, which plays an important role in increasing the control freedom. The geniohyoid showed no significant contribution to any control point. Note that this study only used two specific points to evaluate the contribution of the muscles. If different observation points were adopted, a dissimilar contribution factor might be obtained for some muscles, e.g., the geniohyoid.

To develop a control method, it is necessary to reinspect the EP vectors based on the contribution and the freedom of model control. All of the EP vectors shown in Fig. 8 have a contribution factor larger than 0.3. In the coordinates of the tongue tip, the GGa and verticalis (V) have a similar vector with smaller amplitude. These two vectors are taken into account because they increase the degree of the freedom. The transversus (T) is also taken into account because it increases the degree of freedom, although its role is not explicit in a 2D representation.

In the dorsal coordinate, the EP vector of the MH is located in the space between GGp and SG and contributes to widening the area of contact between the dorsum and the palate, which is required for generating /k/ gestures with different contexts. For this reason, the MH is included in the dorsal coordinate for the control, although its vector is smaller than that of the extrinsic muscles. The GGa and V have a similar EP vector to that of the GGm, but their amplitude is about half that of the GGm. These muscles are not considered an independent factor in the control method because they neither contribute significantly nor increase the degree of freedom. For the same reason, the SL is not treated as an independent factor. However, these muscles are taken into account in the muscle co-contraction, described in the subsequent section.

Based on the above consideration, five major muscles are taken into account in controlling the tongue dorsum, and nine muscles and one muscle group are used for the tongue tip. Thus, the mapping between the spatial points and the muscle forces can be obtained based on the selected EP vectors. An example is shown in the right panel of Fig. 8 by a contour net, which consists of the EPs of the SG and HG. The contour lines correspond to the six force levels. Such a net of contour lines is named the *equilibrium position map* (EP map). With the EP map, any arbitrary point inside the region of the map can be reached using the forces interpolated from the contour lines. The primary difference between the EP map and the EPH (λ model) is that the EP map is a straightforward mapping between muscle forces and the equilibrium positions of the articulators, while the EPH not only requires a muscle loading function but also the λ commands that are involved in a hypothesis on the threshold of the stretch reflex.

3. EP map of the tongue vs jaw positions

Since there is a high correlation between the movements of the tongue and jaw, the coordinate system consisting of the EPs depends strongly on the jaw position. To examine the relationship between the EP maps and jaw positions, we first investigated the equilibrium position of the jaw complex. Five forces of 0.5, 1.2, 2.5, 4.0, and 6.0 N with a duration of 300 ms were applied to the jaw-opener muscle individually, and three forces of 0.5, 1.5, and 3.0 N were applied to the jaw-closer muscle. Corresponding to the forces on the jaw opener, the jaw achieved EPs of -0.16, -0.37, -0.76, -1.10, and -1.42 cm, where the minus sign denotes a position lower than the rest position. When the jaw closer was activated, the jaw reached EPs of 0.09, 0.21, and 0.47 cm in the close direction. Altogether, nine equilibrium positions were obtained for the jaw, which includes one rest position, five opener positions, and three closer positions. The EP maps for the tongue were constructed corresponding to the nine positions of the jaw.

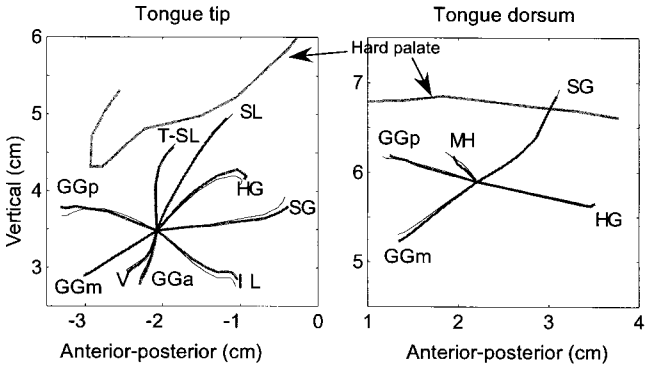


FIG. 9. Example of using a dynamic EP map (the thin lines) derived from the typical EP map of the jaw opener at 0.5 N to represent the EP map (the thick lines) of the jaw opener at 2.5 N.

Comparison of the nine jaw-dependent EP maps shows that the entire EP map rotates and translates with the jaw movement, but the detailed structures of the EP maps has no definite changes. This suggests that it is possible to derive a *dynamic* EP map for any arbitrary jaw position from one EP map by means of translation and rotation. To obtain an optimal dynamic EP map for all the jaw positions, the EP map generated by applying 0.5 N on the jaw-opener muscle is chosen as the typical one, since it is the neutral one in the nine positions. Figure 9 shows an example using a dynamic EP map (the thin lines) derived from the typical EP map to represent the EP map (the thick lines) of the jaw opener with 2.5 N. The rotation degree for the dorsal coordinate was the same as that of the jaw, while a slight inverse rotation was carried out for the tongue tip. In this representation, the average difference between the two EP maps was 0.049 cm for the tongue tip and 0.041 cm for the dorsum.

The accuracy of representing the jaw-dependent EP maps by the typical EP map was evaluated by a mean-squared error between the derived dynamic EP maps and the original ones. Table IV shows the evaluation for all cases, where the top corresponds to the closed positions of the jaw, and the bottom to the wide-open jaw position. The average error over all jaw positions is 0.043 cm for the tongue tip and 0.034 cm for the dorsum. The error increases as the jaw deviates from the reference position. The largest error of 0.12 cm occurred for the tongue tip when the jaw opened about 1.8 cm. For almost all articulations, the tongue tip is not so

TABLE IV. Differences between derived and original EP maps.

Status of the jaw	Force on the jaw muscle (N)	Position on the initial (cm)	Errors for the apex (cm)	Errors for the dorsum (cm)
Closer	3.0	0.407	0.037	0.030
Closer	1.5	0.210	0.026	0.024
Closer	0.5	0.090	0.021	0.020
Initial	0.0	0.00	0.024	0.019
Opener[a]	0.5	−0.160[b]	0.000	0.000
Opener	1.2	−0.374	0.021	0.018
Opener	2.5	−0.761	0.049	0.041
Opener	4.0	−1.097	0.084	0.066
Opener	6.0	−1.425	0.123	0.091

[a]Indicates the typical EP map.
[b]The minus denotes jaw positions lower than the rest position.

crucial when the jaw opens wide. In other words, this representation does not introduce any significant error for crucial points. Figure 9 and Table IV illustrate that a dynamic EP map can be obtained effectively for any jaw position via this derivation.

V. CO-CONTRACTION OF MUSCLES BASED ON THE EP MAP

The most effective way to form a tongue shape by muscle contraction, in view of the minimal energy principle, is for two agonist muscles to work together to achieve a given target. Thus, the muscle forces can be easily estimated via the EP map shown in Fig. 8. During speech, however, the situation is much more complicated because more than two agonist muscles can work together to reach a target, and some agonist–antagonist muscles can co-contract at the same time. To simulate this situation, we designed 17 two-muscle groups and seven three-muscle groups for the tongue tip and five two-muscle groups and one three-muscle group for the dorsum. An example of the EP map for the two-muscle group is shown in the right panel of Fig. 8. The three-muscle group consists of an independent muscle and a muscle pair, in which the activation of the independent muscle corresponds to the co-contraction level and governs a *main* part of the tongue, while the muscle pair manipulates the other part via the mechanism of co-contraction of the agonist and antagonist muscles.

A. Co-contraction between agonist and antagonist

Eight three-muscle groups were designed to generate some potential co-contractions during speech. Figure 10 shows the co-contractions for two three-muscle groups. The thick dark lines show a part of the EP vectors of the coordinates, which were generated by activating the muscles individually. The thin dark lines and the thin gray lines denote the EP trajectories for two synergistic muscles in the group, respectively. The attachment of the thin lines on the EP vector of the independent muscle, indicated by the open circles, corresponds to activation levels of the independent muscle. As shown in the upper panels, the combination of the SG and the muscle group of the GGp and SL can move the dorsum toward the palatal target by the SG and at the same time control the tongue tip to an apical target by the muscle pair. This mechanism can reach a compatible target set for both

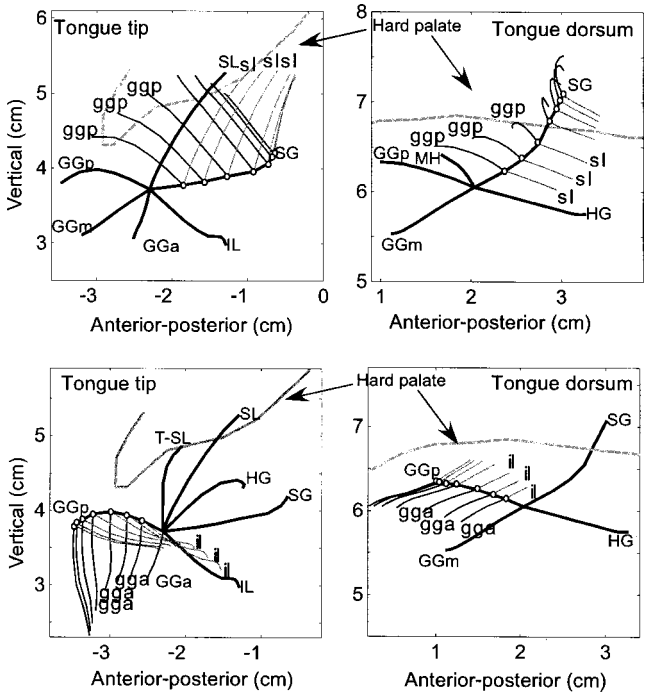


FIG. 10. Co-contractions of the SG and a muscle pair of the GGp and SL (upper panels), and of the GGp and a muscle pair of the GGa and IL (lower panels).

the tongue tip and the dorsum. It is interesting to find that the muscle pair works in synergy for the tongue tip while it functions as an antagonist pair for the tongue dorsum. If the proper force ratio is chosen for the muscle pair of the GGp and SL, the dorsum can be kept in a given position when the tongue tip position is manipulated, and vice versa.

The lower panels in Fig. 10 show the co-contraction between the GGp and the muscle pair of the GGa and IL. This muscle group drives the dorsum to move in a front-upper direction and governs the tongue tip to go downward and backward. The muscle pair demonstrates the same function as that in the upper panel, in which the muscle pair shows a synergetic function for the tongue tip while it behaves as an antagonist pair for the tongue dorsum. Using this combination, the tongue tip can be retracted without interference to the dorsum because the GGa and IL work as antagonists for the dorsum.

Figure 11 shows other examples of the co-contractions. In the upper panels, the combination of the GGm and a muscle pair of the SL and T-SL demonstrates a function that produces a posture with a lower dorsum and a higher tongue tip. The tongue tip is governed by the muscle pair of two intrinsic muscles, SL and T-SL, and an optimal dorsum position can be maintained by using a proper force ratio for the muscle pair. Compared to the other cases, the intrinsic muscles have a smaller effect on the dorsum. It can be seen that the intrinsic muscle pair also definitely works as an antagonist pair for the dorsum. The lower panels of the figure show a different combination, the T-SL and a muscle pair of the GGm and HG, in which the independent muscle is an intrinsic muscle, and the muscle pair consists of two extrinsic muscles. In contrast to the other groups, this group first governs the tongue tip by the intrinsic muscle and then ma-

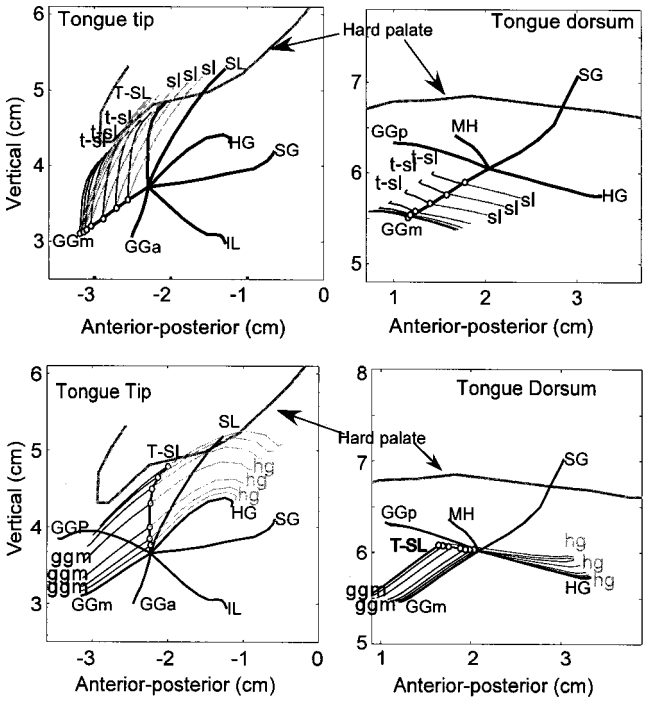


FIG. 11. Co-contractions of the GGm and a muscle pair of the SL and T-SL (upper panels), and the T-SL and a muscle pair of the GGm and HG (lower panels).

nipulates the dorsum by the extrinsic muscles, where the two extrinsic muscles function as an antagonist muscle pair for the tongue tip.

In all of the three-muscle groups, the muscle pair works in synergy for one part of the tongue while behaving as an antagonist pair for the other part. This function can increase the degree of the freedom for model control, allowing us to control different parts of the tongue independently to some extent. Because of this property, we might say that the co-contraction can also be used to maintain the stability of a kinematic system when part of the system is manipulated.

B. Realization of multiple features using co-contraction mechanism

Generally, in a target vector, only one crucial feature is decisive in forming the phoneme, where the other features are referred to as *indecisive* features. To produce a planned sound, the accuracy of the crucial feature must be guaranteed, while an optimal achievement is also desired for the indecisive features. This requires us to deliberate over more than one feature at the same time to find a force set that is optimal for the concerned target features. Figure 12 shows an example of realizing this procedure using a three-muscle combination consisting of the HG and a muscle pair of GGp and SL. In this example, we estimate muscle forces for a given consonantal target with two features for the tongue tip and dorsum, respectively, where the feature for the jaw position was taken into account in generating the jaw-dependent EP map. The crucial feature for the tongue tip is shown by the filled circle, and an indecisive feature for the dorsum is indicated by the larger open circle, which accounts for the coarticulatory effects of the surrounding vowels.

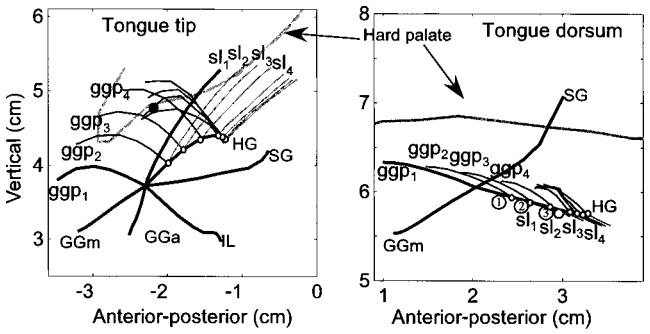


FIG. 12. Setting the apical and dorsal targets simultaneously by means of the co-contraction of the HG and a muscle pair of the GGp and SL. The circles denote the targets for the control points. The circled numbers indicate the candidate targets corresponding to different co-contraction levels.

In Fig. 12, one can see that the HG is the major muscle driving the tongue dorsum toward the dorsal target, while the muscle pair of the GGp and SL is one of the muscle groups that can drive the tongue tip to reach the apical target. Among the force combinations, for example, three combinations of the GGP_1-SL_1, GGP_2-SL_2 and GGP_3-SL_3 are capable of moving the tongue tip to reach the apical target. The difference between them is that they correspond to different co-contraction levels of the HG, whose forces were 0.0, 0.1, and 0.2 N. Since all three combinations can guarantee the crucial feature, the decision of the force set finally depends on their behavior on the indecisive feature. The circled numbers in the dorsal coordinate are the predicted locations for these three force sets. The location of ③ is the closest one to the given dorsal target among the three sets. Therefore, the force set of the GGP_3-SL_3 and HG with 0.2 N is the optimal one for the given target with two features.

In a general process for estimating activation patterns, all possible force combinations are searched out through the EP maps for a given target. The square summation of the muscle forces is calculated for each combination, and the distance between the given target and the candidate target is computed for the indecisive features. The cost function for determining the force set is the weighted summation of the square sum of the forces and the distance. The principle for determining the force set is that it should accurately guarantee the primary feature and optimally realize the indecisive features.

VI. SUMMARY AND DISCUSSION

This study consists of three parts: (1) improve the modeling of tongue tissue; (2) investigate the muscle functions using model simulation; and (3) develop an estimation method for muscle activation forces from spatial targets while considering the co-contraction of muscles.

A. Improvement of the model

The previous version of our model (Dang and Honda 2001; 2002) used volumeless springs as a mesh component, which are replaced by viscoelastic cylinders in the current version. The cylinder is a continuum that fully obeys the physical laws described by the Young's modulus and the Poisson ratio. After dividing the tongue tissue into a hexahe-

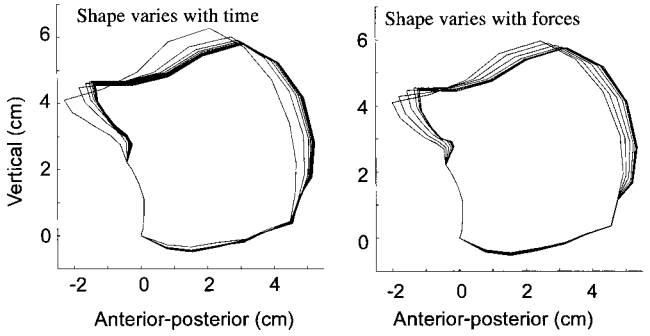


FIG. 13. Temporal deformation of the tongue when a force of 1.5 N is applied on HG (left panel), and the tongue shapes when the HG reaches the equilibrium position for given forces from 0.1 to 4.0 N (right panel).

dron assemblage, each hexahedron was filled with 28 cylinders connecting its vertices in all possible directions. This assemblage of cylinders provides an efficient way to describe the deformations of a hexahedron, because the force applied on any node is decomposed and shared by all connecting cylinders. The continuum properties were achieved to some extent by using the Poisson ratio in the volume constraints of Eq. (6). However, the truss of the cylinders is not a continuum but is discrete. This may introduce some differences from the standard FEM. The result obtained from the test cuboid showed that the semicontinuum model is sufficient for modeling the continuum body for a first-order approximation.

The major benefit of using the semicontinuum model is that a fast computation can be achieved without any significant loss in accuracy. Computation times were about 40 times the real time for simulating articulatory movements, which are more than two orders of magnitude faster than that using the standard FEM. Another advantage is that the truss–frame structure is appropriate for large deformations such as tongue movements. No divergences occurred even when a very large force was applied on the proposed model or an extreme deformation took place during articulatory movements.

B. Behavior of the muscles

This study demonstrated a systematic simulation for all of the tongue muscles. The muscles' contribution to tongue movements was evaluated by the normalized amplitude of the EP vectors. Most extrinsic muscles (except the GGa) govern both the tongue tip and tongue dorsum. The SL and IL showed clear contributions to the tongue tip movements.

The simulation revealed some differences between the results of the present study and those of other studies. First, the HG moves the tongue tip backward and slightly upward but not downward. This is somewhat different from the commonly held view that the HG may also lower the tongue tip. Figure 13 shows some simulations to clarify the causes. The left panel demonstrates the temporal deformation of the tongue when a force of 1.5 N is applied on the HG, and the right panel shows the tongue shapes when the HG reaches an equilibrium position for given forces from 0.1 to 4.0 N. The simulation shows that the tongue moves backward and downward with a slight rotational motion under both conditions. This rotation might explain why the HG does not produce a downward movement of the tongue tip. In tagged MRI observations, Davis *et al.* (1996) found that a local depression occurred behind the tongue tip for steady-state /a/ and suggested that it is due to a shear deformation of the HG with GGa contraction. Analysis of the tagged cine-MRI supported the idea that the GGa produces a tongue-lowering motion during the production of /a/ (Stone *et al.*, 2001). It seems that when the tongue tip moves downward with the tongue dorsum the GGa contracts, assisting the HG.

For the function of the GGp, previous observations showed that the GGp moves the tongue dorsum first forward and then upward (Baer *et al.*, 1988; Alfonso *et al.*, 1987). Our model simulation could not confirm this phenomenon. As shown in Fig. 7, the temporal trajectory did not show any definite tendency in the temporal order of the upward and forward movements resulting from the GGp. The simulation showed that the trajectories depend almost entirely on the force level and the history of the deformation.

A few previous studies stated that the SL, in combination with the GGa, could elevate the tongue tip (Stone *et al.*, 2001; Napadow *et al.*, 1999). This was based on two observations: GGa fibers do not curve forward or extend into the tongue tip, and GGa contraction locally stiffens the tongue behind the tongue tip. Our simulation showed that the SL moves the tongue tip upward, while the GGa has an antagonistic function with the SL to move the tongue tip as shown in Fig. 8. From the point of view of the minimum energy principle or co-contraction, it seems difficult to find a reason for the use of these two muscles as a pair to elevate the tongue tip. However, we cannot negate the suggestion by other studies since the tongue blade was not observed in the current study.

C. Mapping from spatial targets to muscle forces

The equilibrium position for each muscle monotonically shifts as the muscle force increases. Based on the equilibrium position, the EP map was proposed as a way to construct a straightforward mapping between the spatial positions and muscle forces. The force estimation based on the EP map focuses on the spatial target, i.e., the ending point, alone. The estimated force activates a muscle constantly until a new target is coming. Such a force actually is a static force, since dynamic movements of the model were not considered during either force estimation processing or the activation period. For certain large deformations, as shown in Fig. 7, the trajectory is not a straight line. It is difficult to define such a curved trajectory by using its starting and ending points only. Actually, the trajectory can also be another measure for the spatial target (Okadome and Honda, 2001). Achieving trajectory control also requires us to estimate the dynamic component of the forces. The authors have developed a method, namely *muscle workspace*, to estimate the dynamic muscle forces stepwise by minimizing the distance from a current position and the target (Dang and Honda, 2002). Combination of both of the static and dynamic force estimations in the model control method remains a development task for future work.

To simulate co-contractions between agonist and antagonist muscles, we designed several three-muscle combinations, consisting of one independent muscle and a pair of two muscles. The simulation showed that the two muscles of the muscle pair worked in synergy for one control point while functioning as an antagonist muscle pair for the other one. This function can be used to keep the stability of a kinematic system when a part of the system is manipulated. This property can be used on the model to achieve multiple targets simultaneously, based on a strategy that accurately guarantees the crucial target and optimally reaches the indecisive targets.

A method was proposed in this study to estimate the muscle patterns for a given target based on the EP maps, in which all possible force combinations are searched out through the EP maps, and then one optimal force set is chosen to accurately guarantee the primary feature and optimally reach the indecisive features. Model simulation showed that this method is an efficient way to achieve multiple targets by using the co-contraction mechanism. However, our proposed method may be far different from that used by the human brain, which is still too complicated for humans to learn and use to optimize muscle activation patterns. Therefore, a more efficient method is still desired.

ACKNOWLEDGMENTS

A part of this study was finished when the first author was working for the Institut de la Communication Parlee (ICP) Grenoble, France. The authors would like to thank Pascal Perrier for general discussions as well as thank Mark Tiede, Donna Erickson, and Pierre Badin for their valuable comments. The authors deeply appreciate their instructive comments of Anders Löfqvist and two anonymous reviewers of this paper. This research has been supported in part by the Telecommunications Advancement Organization of Japan.

APPENDIX: ANALYSIS OF DISPLACEMENT-BASED FINITE ELEMENT METHOD

The present study adopts a displacement-based finite-element method as the basis of the modeling, which is referred to as an extended finite element method (X-FEM). The principal advantages of X-FEM are that the finite-element framework (sparsity and symmetry of the stiffness matrix) is retained and that a single-field variational principle is used (Zienkiewicz and Taylor, 1989; Belytschko *et al.*, 2002).

To apply the X-FEM to our modeling, the equilibrium of a general 3D body is first considered, as shown in Fig. 14. The body is located in the fixed coordinate system *X, Y, Z*. The body surface area is supported on the area $\mathbf{S}_u$ with prescribed displacements $\mathbf{U}^{Su}$ and is subjected to surface traction $\mathbf{f}^{Sf}$ on the surface area $\mathbf{S}_f$. In addition, the body is subjected to the externally applied body force $\mathbf{f}^B$ and concentrated loads R^i (where i denotes the point of load application). The displacements of the body from the unloaded configuration are measured in the equation

$$P(X,Y,Z)=[U,V,W]^T, \tag{A1}$$

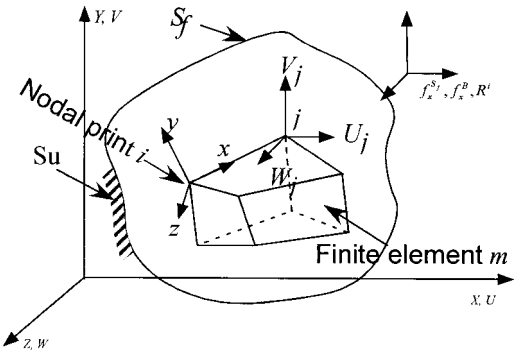


FIG. 14. General three-dimensional description of the displacement-based finite-element method.

and $P=P^{Su}$ on the surface area Su. The strains corresponding to P are

$$\varepsilon=[\varepsilon_{XX}\varepsilon_{YY}\varepsilon_{ZZ}\varepsilon_{XY}\varepsilon_{YZ}\varepsilon_{ZX}], \tag{A2}$$

where $\varepsilon_{XX}=\partial P/\partial X$, $\varepsilon_{YY}=\partial P/\partial Y$, $\varepsilon_{ZZ}=\partial P/\partial Z$, $\varepsilon_{XY}=\partial P/\partial X+\partial P/\partial Y$, $\varepsilon_{YZ}=\partial P/\partial Y+\partial P/\partial Z$, $\varepsilon_{ZX}=\partial P/\partial Z+\partial P/\partial X$. The stresses corresponding to ε are

$$\tau=[\tau_{XX}\tau_{YY}\tau_{ZZ}\tau_{XY}\tau_{YZ}\tau_{ZX}], \tag{A3}$$

where $\tau=C\varepsilon+\tau^I$. C is the stress–strain material matrix and the vector τ^I denotes given initial stresses.

The basis of the displacement-based finite-element solution is the principle of virtual displacements, i.e., the principle of virtual work. This principle states that the equilibrium of the body in Fig. 14 requires that for any compatible small virtual displacements imposed on the body in its state of equilibrium, the total internal virtual work is equal to the total external virtual work

$$\int_V \varepsilon^T\tau\,dV=\int_V P^T f^B\,dV+\int_{s_f} P^{s_f T} f^{s_f} dS+\sum_i P^{iT}R^i, \tag{A4}$$

where P is the virtual displacement and ε is the corresponding virtual strain. When the principle of virtual displacements is satisfied for all virtual displacements with stresses τ, all three fundamental requirements of the mechanics, equilibrium, compatibility, and the stress–strain law are fulfilled. Stresses τ are obtained from a continuous displacement field P that satisfies the displacement boundary conditions on $\mathbf{S_u}$.

In the finite-element analysis, we approximate the body in Fig. 14 as an assemblage of discrete finite elements interconnected at the nodal points on the element boundaries. Displacement p measured in a local coordinate system *x,y,z* within each element is assumed to be a function of the displacement at all nodal points bounded on the element. Therefore, for element m we have

$$p^{(m)}(x,y,z)=H^{(m)}(x,y,z)X^{(m)}, \tag{A5}$$

where $H^{(m)}$ is the displacement interpolation matrix and $X^{(m)}$ is a vector of global displacement components in three dimensions for all nodal points.

Figure 14 shows a typical finite element of the assemblage. This element has eight nodal points, one at each corner, which can be thought of as a ''brick'' element. The complete body is represented as an assemblage of such brick

elements put together so that no gap is left between the element domains. Therefore, for a given brick element, the displacement at the nodes can fully describe the displacement and strain distributions within the element.

With the assumption on the displacements in the above equation, the corresponding element strain is given by

$$\varepsilon^{(m)}(x,y,z) = D^{(m)}(x,y,z)X^{(m)},$$
$$\tau^{(m)}(x,y,z) = C^{(m)}\varepsilon^{(m)}(x,y,z) + \tau^{I(m)}, \tag{A6}$$

where $D^{(m)}$ is the strain-displacement matrix, and the rows of $D^{(m)}$ are obtained by appropriately differentiating and combining the rows of the $H^{(m)}$. $C^{(m)}$ is the stress–strain matrix of the material, which depends only on the Young's modulus and Poisson ratio. An effective assembly process of all matrices into the governing structure matrices can be achieved by using the above equations in the principle of virtual displacement. This process is referred to as *the direct stiffness method*.

Using the direct stiffness method, the stiffness matrix K can be calculated for the element assemblage

$$K = \sum_m \int_{V^{(m)}} D^{(m)^T} C D^{(m)} dV^{(m)}, \tag{A7}$$

where the generalized stress–strain matrix C becomes independent of the elements with the assumption that the material is isotropic. Using d'Alembert's principle, the element inertia forces can be included in the body forces. Thus, the mass matrix M of the structure can be obtained from the integral of (A8). For the velocity-dependent damping forces, the damping matrix B can be obtained in the same way

$$M = \sum_m \int_{V^{(m)}} \rho H^{(m)^T} H^{(m)} dV^{(m)},$$
$$B = \sum_m \int_{V^{(m)}} b H^{(m)^T} H^{(m)} dV^{(m)}, \tag{A8}$$

where ρ is the mass density, and b is the damping property parameter.

Based on the description above, the derived equation of equilibrium governing the linear dynamic response of a finite element system is

$$M\ddot{X} + B\dot{X} + KX = F, \tag{A9}$$

where X, $\dot{X}$, and $\ddot{X}$ are the displacement, velocity, and acceleration vectors of the finite element assemblage, respectively. F is the vector of externally applied loads, described in the following integral:

$$F = \sum_m \left(\int_{V^{(m)}} H^{(m)} f^B \, dV^{(m)} + \int_{S^{(m)}} H^{S(m)} f^S \, dS^{(m)} \right) + f^L, \tag{A10}$$

where f^B, f^S, and f^L represent body, surface, and concentrated forces, respectively.

According to X-FEM, the matrices M, B, and K, and the force F are derived from (A7) and (A8), in which the description is based on the displacements of the nodal points. For a loading force, thus, the displacements X can be obtained from (A9), and then the stress can be evaluated using (A6).

Alfonso, P. J., Watson, B. C., and Baer, T. (**1987**). ''Measuring stutterers' dynamical vocal tract characteristics by x-ray microbeam pellet tracking,'' in *Speech Motor Dynamics in Stuttering*, edited by H. F. M. Peters and W. Hulstijn (Springer, New York).

Baer, T., Alfonso, J., and Honda, K. (**1988**). ''Electromyography of the tongue muscle during vowels in /∂pvp/ environment,'' Ann. Bull. R. I. L. P., Univ. Tokyo **7**, 7–18.

Bathe, K. (**1996**). *Finite Element Procedures* (Prentice-Hall, Englewood Cliffs, NJ).

Belytschko, T., Parimi, C., Moës, N., Sukumar, N., and Usui, S. (**2002**). ''Structured extended finite element methods for solids defined by implicit surfaces,'' Int. J. Numer. Methods Eng. **56**(4), 609–635.

Dang, J., and Honda, K. (**1997**). ''Correspondence between three-dimensional deformation and EMG signals of the tongue,'' Proceedings of the Spring Meeting of Acoust. Soc. Jpn., 241–242 (in Japanese).

Dang, J., and Honda, K. (**1998**). ''Speech production of vowel sequences using a physiological articulatory model,'' Proc. ICSLP98, 5, 1767–1770.

Dang, J., and Honda, K. (**2001**). ''A physiological model of a dynamic vocal tract for speech production,'' Acoust. Sic. Tech **22**, 415–425.

Dang, J., and Honda, K. (**2002**). ''Estimation of vocal tract shape from sounds via a physiological articulatory model,'' J. Phonetics **30**, 511–532.

Dang, J., Honda, M., and Honda, K. (**2002**). ''Investigation of coarticulation based on electromagnetic articulographic data,'' Proc. ICSLP2002, 1521–1524.

Dang, J., Honda, K., and Tohkura, Y. (**1997**). ''3D observation of the tongue articulatory movement for Chinese vowels,'' Technical Report of IEICE, SP97–11, 9–16.

Dang, J., Sun, J., Deng, L., and Honda, K. (**1999**). ''Speech synthesis using a physiological articulatory model with feature-based rules,'' Proc. ICPhS-99, 2267–2270.

Davis, E., Douglas, A., and Stone, M. (**1996**). ''A continuum mechanics representation of tongue motion in speech,'' Proceedings of ICSLP1996, 788–792.

Duck (**1990**). *Physical Property of Tissues: A Comprehensive Reference Book* (Academic, London).

Erickson, D. (**2002**). ''Articulation of extreme formant patterns for emphasized vowels,'' Phonetica **59**, 134–149.

Feldman, A. (**1986**). ''Once more on the equilibrium-point hypothesis (λ model) for motor control,'' J. Motor Behav. **1**, 17–54.

Fung, Y. (**1993**). *Biomechanics—Mechanical Properties of Living Tissue*, 2nd ed. (Spriger, New York).

Hashi, M., Westbury, J., and Honda, K. (**1998**). ''Vowel posture normalization,'' J. Acoust. Soc. Am. **104**, 2426–2437.

Hashimoto, K., and Suga, S. (**1986**). ''Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue,'' J. Acoust. Soc. Jpn. (E) **7**, 39–46.

Hill, A. V. (**1938**). ''The heat of shortening and the dynamic constants of muscle,'' Proc. R. Soc. London, Ser. B **126**, 136–195.

Hirai, H., Dang, J., and Honda, K. (**1995**). ''A physiological model of speech organs incorporating tongue–larynx interaction,'' J. Acoust. Soc. Jpn. **52**, 918–928 (in Japanese).

Honda, K., Hirai, H., and Dang, J. (**1994**). ''A physiological model of speech organs and the implications of the tongue-larynx interaction,'' Proc. ICSLP94, 175–178, Yokohama.

Kakita, Y., Fujimura, O., and Honda, K. (**1985**). ''Computational of mapping from the muscular contraction pattern to formant pattern in vowel space,'' in *Phonetic Linguistics*, edited by A. L. Fromkin (Academic, New York).

Kiritani, S., Miyawaki, K., Fujimura, O., and Miller, J. (**1976**). ''A computational model of the tongue,'' Ann. Bull. R. I. L. P. Univ. Tokyo **10**, 243–251.

Laboissière, R., Ostry, D., and Feldman, A. (**1996**). ''The control of multi-muscle system: Human jaw and hyoid movement,'' Biol. Cybern. **74**, 373–384.

Min, Y., Titze, I., and Alipour, F. (**1994**). ''Stress–strain response of the human vocal ligament,'' NCVS Status Prog. Rep. **7**, 131–137.

Miyawaki, K. (**1974**). ''A study of the muscular of the human tongue,'' Ann. Bull. R. I. L. P. Univ. Tokyo **8**, 23–50.

Morecki, A. (**1987**). ''Modeling, mechanical description, measurements and control of the selected animal and human body manipulation and locomo-

tion movement,'' in *Biomechanics of Engineering—Modeling, Simulation, Control*, edited by A. Morecki (Springer, New York), pp. 1–28.
Napadow, V. J., Chen, Q., Wedeen, V. J., and Gilbert, R. J. (**1999**). ''Intramural mechanics of the human tongue in association with physiological deformations,'' J. Biomech. **32**, 1–12.
Niimi, S., Kumada, M., and Niitsu, M. (**1994**). ''Functions of tongue-related muscles during production of the five Japanese vowels,'' Ann. Bull. R. I. L. P. Univ. Tokyo **28**, 33–40.
Oka, S. (**1974**). *Rheology–Biorheology* (Shokabo, Tokyo, 1974), pp. 445–456.
Okadome, T., and Honda, M. (**2001**). ''Generation of articulatory movements by using a kinematic triphone model,'' J. Acoust. Soc. Am. **109**, 453–463.
Payan, Y., and Perrier, P. (**1997**). ''Synthesis of V–V sequences with a 2D biomechanical tongue shape in vowel production,'' Speech Commun. **22**, 185–206.
Perkell, J. (**1974**). ''A physiological-oriented model of tongue activity in speech production,'' Ph.D. thesis, MIT.
Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., and Jackson, M. (**1992**). ''Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements,'' J. Acoust. Soc. Am. **92**, 3078–3096.
Sanguineti, V., Laboissière, J., and Payan, Y. (**1997**). ''A control model of human tongue movements in speech,'' Biol. Cybern. **77**, 11–22.
Sanguineti, V., Laboissière, J., and Ostry, D. (**1998**). ''A dynamic biomechanical model for neural control of speech production,'' J. Acoust. Soc. Am. **103**, 1615–1627.
Stevens, K. (**2000**). *Acoustic Phonetics* (The MIT Press, Cambridge, MA).
Stone, M., Davis, E., Douglas, A., Ness Aiver, M., Gullapalli, R., Levine, W., and Lundberg, A. (**2001**). ''Modeling motion of the internal tongue from tagged cine-MRI images,'' J. Acoust. Soc. Am. **109**, 2974–2982.
Takemoto, H. (**2001**). ''Morphological analyses of the human tongue musculature for three-dimensional modeling,'' J. Speech Lang. Hear. Res. **44**, 95–107.
Warfel, J. (**1993**). *The Head, Neck, and Trunk* (Led & Febiger, Philadelphia and London).
Wilhelms-Tricarico, R. (**1995**). ''Physiological modeling of speech production: Methods for modeling soft-tissue articulators,'' J. Acoust. Soc. Am. **97**, 3805–3898.
Yamazaki, K. (**1933**). ''The weight of the cranium and mandible with comparison of the dental and bony regions of the mandible,'' Jpn. J. Dentistry **26**, 769–796 (in Japanese).
Zajac, F. (**1989**). ''Muscle and tendon: Properties, models, scaling, and application to biomechanics and motor control,'' Crit. Rev. Biomech. Eng. **17**, 359–411.
Zienkiewicz, O., and Taylor, R. (**1989**). *The Finite Element Method* (McGraw-Hill, New York).

Development of [ɹ] in young, Midwestern, American children

Richard S. McGowan
CReSS LLC, 1 Seaborn Place, Lexington, Massachusetts 02420-2002

Susan Nittrouer
Center for Persons with Disabilities, Utah State University, Logan, Utah 84322-6885

Carol J. Manning
Boys Town National Research Hospital, Omaha, Nebraska 68131

(Received 12 September 2002; revised 29 October 2003; accepted 24 November 2003)

Beginning at the age of about 14 months, eight children who lived in a rhotic dialect region of the United States were recorded approximately every 2 months interacting with their parents. All were recorded until at least the age of 26 months, and some until the age of 31 months. Acoustic analyses of speech samples indicated that these young children acquired [ɹ] production ability at different ages for [ɹ]'s in different syllable positions. The children, as a group, had started to produce postvocalic and syllabic [ɹ] in an adult-like manner by the end of the recording sessions, but were not yet showing evidence of having acquired prevocalic [ɹ]. Articulatory limitations of young children are posited as a cause for the difference in development of [ɹ] according to syllable position. Specifically, it is speculated that adult-like prevocalic [ɹ] production requires two lingual constrictions: one in the mouth, and the other in the pharynx, while postvocalic and syllabic [ɹ] requires only one oral constriction. Two lingual constrictions could be difficult for young children to produce. [DOI: 10.1121/1.1642624]

PACS numbers: 43.70.Ep [AL] Pages: 871–884

I. INTRODUCTION

A. Phonetics and acoustics of /r/

In rhotic dialects of American English, the /r/ phoneme is pronounced as an approximant [ɹ], and it is notoriously difficult for American children to learn to produce.[1] Sander (1972) reported that the median age for acquisition of /r/ for American children was 3 years, and it was not until age 6 years that 90% of children produced /r/ correctly. Smit *et al.* (1990), in their study of 3- to 9-year-old children from Iowa and Nebraska, reported that 90% of the children had attained correct /r/ production by 8 years of age.

In adults' productions of prevocalic [ɹ], the most clearly defining acoustic property is a very low third formant frequency, $F3$. It often dips below 2000 Hz, which is well below its value for a neutral vowel. Another acoustic correlate of [ɹ] is that $F3$ is generally close to $F2$. That is, the value of $F3-F2$ is smaller in [ɹ] than it is for a neutral vowel (Lehiste, 1962; Delattre and Freeman, 1968). However, variation exists in the precise acoustic properties of [ɹ] depending on whether it is prevocalic, postvocalic, or syllabic (Lehiste, 1962; Delattre and Freeman, 1968; Olive, Greenwood, and Coleman, 1993). The word ''right'' contains a prevocalic [ɹ]; the word ''car'' possesses a postvocalic [ɹ]; the name ''Burt'' has a medial syllabic [ɹ]; and the word ''doctor'' has a final syllabic [ɹ]. In general, the first three formant frequencies are not as low in postvocalic and syllabic [ɹ] relative to their values in a neutral vowel, as they are in prevocalic [ɹ] in a stressed syllable. However, $F3$ is lower and $F3-F2$ smaller than their expected values in neutral vowels for all of these variations of [ɹ]. The properties of low $F3$ and smaller $F3-F2$ are what we call properties of [ɹ]-*ness*. Thus, the previous literature indicates that details of the formant structure for an [ɹ] depends on its syllable position.

The literature on young American children's production of /r/ is sparse and often limited to word-initial or prevocalic /r/. Dalston (1975) studied the formant frequencies of word initial /r/, /w/, and /l/ in adults and children who were 3 to 4 years of age. He confirmed that a relatively low $F3$ occurred when the children produced word-initial /r/ correctly as [ɹ], and that this $F3$, with a mean of approximately 2500 Hz, helped to distinguish /r/ from /w/ and /l/, which both had mean $F3$'s of approximately 3500 Hz. Also, scatter plots of $F3/F1$ versus $F2/F1$ for the adults and children revealed that the children produced /r/ and /w/ with more overlap in the $F2/F1$ parameter than did the adults. This overlap could be a contributing factor to adults perceiving /w/ when word-initial, prevocalic /r/ is produced incorrectly by children (e.g., Smit *et al.*, 1990; Smit, 1993; Shriberg and Kent, 1995). In a phonetic study of segment acquisition in children 15 to 24 months of age, Stoel-Gammon (1985) found that word-final /r/ was acquired before word-initial /r/. This corroborated other, earlier studies of American English speech sound acquisition. The phonetic studies of Smit *et al.* (1990) and Smit (1993) indicated that 2- to 5-year-old children from Iowa and Nebraska could produce intersyllabic /r/ and syllabic /r/ with substantially fewer errors than they could produce prevocalic /r/.

The protracted period of development of [ɹ] allows us to examine the development of an English speech segment. Further, the acoustic data on adult [ɹ] indicate that the way it is produced depends on syllable position. We were interested in whether young children also show differences in production of this segment depending on syllable position, as well as whether the speed with which [ɹ] developed depended on syllable position. While similar developmental trends may

exist for segments other than [ɹ], they might be more easily missed in longitudinal studies if children progress through the stages of development rapidly.

B. Articulatory considerations and articulatory–acoustic relations for /r/

The acoustic correlates of the [ɹ] segment can be described, and it is known that several different articulatory gestures can give rise to these acoustic characteristics. Yet, an aspect that provides an understanding of the development of [ɹ] is an understanding of the articulatory gestures that children use to produce /r/ and how these gestures differ across syllable position. Along with a review of the literature of adult's [ɹ] production, an example of simultaneously recorded acoustic and articulatory data from an adult talker will be examined below to develop a hypothesis about the possible articulatory gestures that could underlie the acoustic data that were collected from children. The procedures required to obtain data on the articulatory gestures involved in /r/ production are more invasive than generally considered acceptable for work with young children. While adults cannot provide much insight into children's articulatory behavior, they can provide insight into the physical relations between articulation and acoustics.

Delattre and Freeman (1968) performed an extensive study of adults' American English [ɹ] using x-ray cineradiography with simultaneous acoustic recording. Their results showed that each of their American English speakers produced stressed, prevocalic [ɹ] with two vocal-tract constrictions: one palato-velar and the other pharyngeal. There was a variety of tongue shapes, with either the tongue dorsum, tongue blade, or tongue tip providing the palato-velar constriction. Also, a variety of constriction degrees was used, and narrower constrictions were correlated with lower $F3$ frequencies (and so with smaller $F3-F2$ differences, as well). Furthermore, lip rounding was used consistently by their speakers for prevocalic /r/ in stressed syllables. This had the effect of lowering $F1$, $F2$, and $F3$ from their values in postvocalic /r/, which was not produced with lip rounding. Otherwise, Delattre and Freeman (1968) observed that the postvocalic [ɹ] had the same tongue shapes as prevocalic [ɹ], but with $F3$ not as low and $F3-F2$ not as small as for stressed, prevocalic [ɹ].

The relation between the acoustics and articulation of [ɹ] has received attention more recently. Stevens (1998) proposed a model for retroflexed [ɹ] production in which the volume under the tongue creates an acoustic side branch that gives rise to a pole–zero pair. The pole would constitute a formant that appears between the formants that are continuous with the $F2$ and $F3$ of the surrounding vowels. However, a different model of [ɹ] production, based on the three-dimensional MRI data of Alwan, Narayanan, and Haker (1997), was proposed by Espy-Wilson *et al.* (2000) and by Jackson, Espy-Wilson, and Boyce (2001). These investigators judged the dimensions of the sublingual cavities for the two subjects producing [ɹ] in the MRI study to be too small to account for an extra formant just above $F2$ and below $F3$, as proposed by Stevens (1998). Rather, they proposed that the cavity in front of the palatal constriction was of sufficient size to produce a low $F3$. Whether this front cavity should be modeled as a Helmholtz resonator or as a quarter-wavelength resonator would depend on the degree of lip rounding. They also showed that the region of the oral cavity behind the tongue tip could be modeled either as a double-Helmholtz resonator or as a single-Helmholtz resonator in series with a half-wavelength resonator accounting for $F1$ and $F2$, depending on the degree of the pharyngeal constriction. Thus, in the case where both the palatal and pharyngeal constrictions are tight and lip rounding is present, the entire system behaves more or less as three coupled Helmholtz resonators, from which three low-frequency formants result. However, if the pharyngeal constriction is only moderate, it is better to model the situation as two coupled Helmholtz resonators, with the coupling at the palatal constriction. The palatal constriction produces a low $F1$ and $F3$, while the moderate pharyngeal constriction lowers $F1$ and $F2$. Espy-Wilson and Boyce (1999) reported that $F4$ is relatively low for retroflexed articulations compared to its value for nonretroflexed articulations. If $F4$ is a resonance of the cavity behind the palatal constriction, then its low value could be the result of the palatal constriction of the retroflexed [ɹ] being farther forward or shorter than for the nonretroflexed [ɹ].

Others have studied American English [ɹ] articulation from different perspectives. For example, using MRI technology Alwan *et al.* (1997) provided much-needed three-dimensional data on sustained [ɹ], from which acoustic tube models could be constructed. One subject, PK, produced both sustained retroflexed and nonretroflexed [ɹ] with both tight palatal and tight pharyngeal constriction. Another subject, MI, produced sustained word-initial and syllabic [ɹ] with a tight palatal constriction, but only a moderate constriction in the pharynx. Lip rounding was involved in all these productions. One of their observations was that American English retroflexed [ɹ] was actually produced by a raised, laminal tongue blade, and not a curled tongue blade. Westbury, Hashi, and Lindstrom (1998) used x-ray microbeam technology with the data from the X-Ray MicroBeam Speech Production Database, XRMB-SPD (Westbury, 1994) to describe a continuum of articulatory shapes from bunched to retroflexed articulation for prevocalic [ɹ] in a large cohort of adult subjects. (A bunched articulation is one in which the tongue body is used to make the palatal constriction.) Guenther *et al.* (1999) used electromagnetic articulometry to show that there was an articulatory trade in seven talkers for stressed, prevocalic [ɹ] production: As the cavity in front of the constriction became shorter, due to articulatory constraints of the phonetic context, the palatal constriction widened and/or constriction length increased, thus enabling $F3$ to remain low in differing phonetic environments. There was also a study of [ɹ] production using the very minimal technology of cotton swabs to find the position of the tongue blade during prevocalic, postvocalic, and syllabic [ɹ] production (Hagiwara, 1994). In this study [ɹ] articulations were classified as tip-up, tip-down, or blade-up, depending on whether a cotton swab through the incisors touched the underside of the tongue blade, the tongue tip, or the upper surface of the blade, respectively. All three articulation types could occur for all syllable positions. The combined result of

these experiments is that speakers use a wide variety of articulatory maneuvers to produce [ɹ], including nonretroflexed tongue bunching near the palate, as well as retroflexed configurations.

The articulatory and acoustic correlates of [ɹ] according to syllable position were of great interest to this work because we wanted to understand the articulatory–acoustic relations children use in producing /r/ in different syllable positions. To further understand the physics of the articulatory–acoustic relations in [ɹ] beyond that provided by the work already cited, we examined a subject from the XRMB-SPD, JW11, who exhibited both a retroflex and bunched prevocalic [ɹ]. In the word ''right'' JW11 produced [ɹ] using a retroflexed tongue blade, while he produced [ɹ] with a nonretroflexed articulation in ''rag.'' Three tokens of the prevocalic [ɹ] in ''right'' and two tokens of prevocalic [ɹ] in ''rag'' were examined. Along with these words containing prevocalic [ɹ], we examined examples of postvocalic [ɹ] in one token of ''there,'' three tokens of ''large,'' and two tokens of ''dormer'' (first syllable), as well as final syllabic [ɹ] in two tokens of ''dormer'' (second syllable). These observations provided the following insights for JW11, the details to which are contained in the Appendix. Given equivalent syllable stress and phrase positions, $F1$, $F2$, and $F3$ were generally higher for postvocalic and syllabic [ɹ] than for prevocalic [ɹ] because: (1) compared to prevocalic [ɹ], there was reduced or no lip rounding for postvocalic and syllabic [ɹ], and (2) the palatal constriction for postvocalic and syllabic [ɹ] was not as tight as for prevocalic [ɹ]. It cannot be determined whether the subject produced varying degrees of pharyngeal constriction across word positions because there is no indication of tongue root or larynx position in the XRMB-SPD. However, we speculate that there was reduced or no pharyngeal constriction for postvocalic and syllabic [ɹ], while there was substantial pharyngeal constriction for prevocalic [ɹ]. In support of this speculation, it was noted that the tokens with the lowest $F2$'s for postvocalic [ɹ] were those with neighboring back vowels, so that any pharyngeal constriction for these [ɹ]s may be the result of carryover coarticulation.

Based on the observations of JW11 and of the previously published work reviewed above, the articulation of [ɹ] according to syllable position can be summarized as follows: Prevocalic [ɹ] is a consonant articulated with at least one close approximation in the palato-velar region, along with a secondary constriction in the pharyngeal region, and some degree of lip rounding. Postvocalic [ɹ] is more of an off-glide to the preceding vowel with one primary constriction target in the palatal region with little or no lip rounding. Syllabic [ɹ] is a monophthong vowel with a steady constriction that is similar to that of the constriction target of the postvocalic [ɹ].

C. The present study

The present study quantifies observations of children's /r/ production in terms of $F2$, $F3$, and their separation, $F3-F2$. Also, the formant frequency trajectories for /r/ in certain vowel contexts will be compared. Statistical analyses will include linear regression so that changes of the formant frequencies and their separation for /r/'s in different syllable position can be quantified as a function of age. For tokens in similar age groups, rank-order statistics will help quantify the differences between each $F2$, $F3$, and $F3-F2$ as functions of /r/ syllable position. While there are some reasons to believe that there could be differences in children's perception of [ɹ] depending on syllable position and status, we focus on articulatory causes for the differences in the development of [ɹ] production. As reviewed above, the prevocalic [ɹ] is articulatorily more complex than the postvocalic [ɹ] and syllabic [ɹ]. Thus, as has been observed phonetically in previous literature (e.g., Smits, 1993), we expect prevocalic [ɹ] not to be as well developed as the [ɹ]'s in other syllable positions for young children.

II. METHOD

A. Subjects

Nine children from eastern Nebraska and western Iowa were recruited for a longitudinal study of speech production, but one child left the area before completing the study. The children recruited were typically developing children. All had normal prenatal histories, normal deliveries, and no special medical conditions. None of the children had a family member with speech, language, or hearing disorders. None of the children had any reported history of otitis media with effusion at the start of the study, and no child was treated for more than one episode while the study was being conducted.

B. Procedures

Recording sessions were started as soon as possible after each child began producing consistent phonetic forms. At the start of the recording sessions all children were about 1 year of age and had vocabularies of fewer than 10 words. Children were recorded approximately every 2 months. However, parents were asked to contact the laboratory staff if they noticed what appeared to be a particularly rapid proliferation of new words, or when they noticed their child was starting to combine words. Recording sessions were discontinued when a child was consistently using sentences of three or more words, with some function words.

Children were recorded in the same sound-treated chamber at each session. Sessions were 20 min long. The child sat in a highchair at a table, with one parent across the table. The same set of toys was available for play at each session, and consisted of such things as small stuffed dolls, foam puzzles, and cloth books. All toys used in these sessions were soft to minimize extraneous noises that might interfere with speech recording. The toys were not chosen to elicit any particular response from the children. Parents were instructed to play with their children, trying to elicit as much language as possible. Also, parents kept a diary of new vocabulary items (at the younger ages) and new sentence structures (at the older ages) that they heard at home.

Recordings were obtained using an AKG C-535EB microphone, a Shure model M268 mixer, and a Nakamichi MR-2 cassette deck. This system provided a flat-frequency response out to 20 kHz. The microphone was suspended roughly 9 in. above the child's mouth. It was suspended rather than table mounted because pilot work showed that children habituated to its presence more rapidly that way.

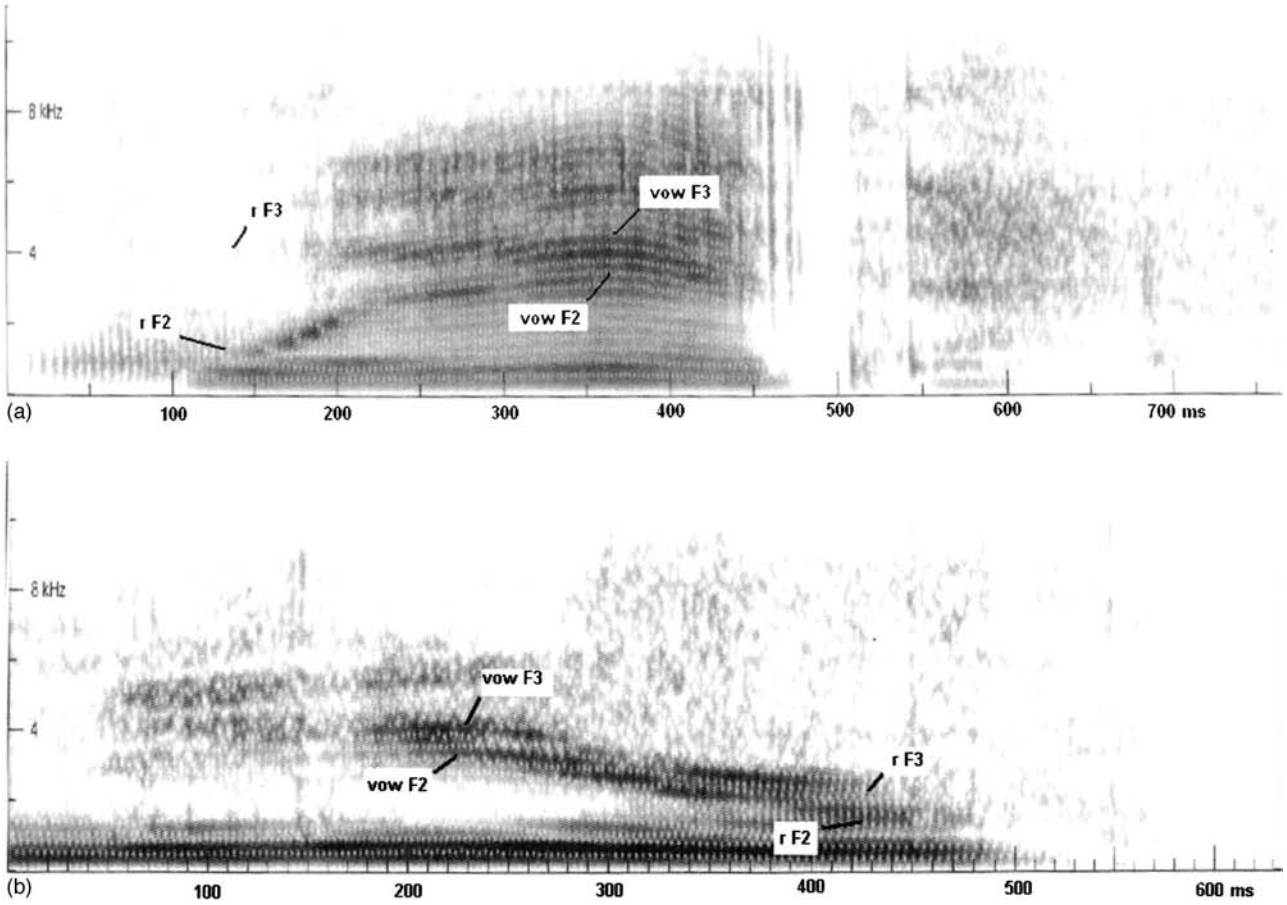


FIG. 1. (a) A spectrogram of a subject's production of the word ''red.'' The $F2$ and $F3$ in the /r/ and in the /ɛ/ are indicated, as measured using spectral cross sections and the spectrogram. (b) A spectrogram of a subject's production of the word ''here.'' The $F2$ and $F3$ in the /r/ and in the /i/ are indicated, as measured using spectral cross sections and the spectrogram.

These recordings were subsequently digitized with a Soundblaster A/D card using SPEECH STATION II software at a 22.05-kHz sampling rate.

The recordings of each child were examined and analyzed. For each child, the final recording was examined first, the penultimate recording second, and so on until the first recording of that child was examined. This allowed the examiner to acclimate to each child's speech starting with what should have been the most intelligible sample. Utterances with words that would have contained an [ɹ] or syllabic [ɹ] if spoken by an adult with a rhotic dialect were extracted. Conversational context helped in the identification of these utterances. Also, the children often repeated a single lexical item while engaged in play. The one exception to the way utterances were selected for analysis was from a child who produced the word ''bakery'' alternately as ''brakery'' or ''bwakery,'' and that inserted /r/ was analyzed. There were no other instances of such /r/ insertions.

Using a spectral analysis program, SPEECH STATION II, the $F2$ and $F3$ of each word were measured in the temporal region most clearly affiliated with the /r/ phoneme, as well as in the middle of the neighboring vowel, using DFT spectral cross sections and spectrograms. Figures 1(a) and (b) show examples of spectrograms of words with prevocalic /r/ (''red'') and postvocalic /r/ (''here''), respectively, as spoken by two of the children. Hamming windows of 256 samples (11.6 ms) were used. The windows overlapped by 5.8 ms, or 128 samples. In the case of prevocalic /r/, the neighboring vowel followed the /r/ (or attempted /r/). For postvocalic /r/, the neighboring vowel preceded the /r/. Intersyllabic /r/ was counted as prevocalic, although there were not many of these. Formant measurements in the diphthongs /aɪ/ and /eɪ/, which only appeared with prevocalic /r/, were taken in the off-glide at the most steady portions and before transition to closing consonants. For prevocalic /r/, $F2$ and $F3$ were measured at the $F2$ minimum. In the case of postvocalic /r/, either the local minimum or local maximum value of $F2$ was chosen as the measurement time, depending on whether the $F2$ trajectory from the preceding vowel was falling or rising. Measurement of medial syllabic /r/, as in ''Burt,'' was done the same way as for postvocalic /r/, except that formants were measured toward the beginning and toward the end of syllabic /r/, while avoiding the surrounding consonantal transitions. This was also the measurement method applied to word-final, unstressed, syllabic /r/, as in ''keeper.'' The two types of syllabic /r/'s, medial and final, were analyzed separately because it was not clear at the outset whether word-final syllabic /r/ would behave more like a vowel-and-postvocalic-/r/ or more like a medial syllabic /r/. Only those words for which $F2$ or $F3$ data could be measured were included in the analysis, so that words which were too faint or produced in a scream were excluded from the analysis.

Further, $F2$ was required to be between 1000 and 3500 Hz and $F3$ between 2000 and 5100 Hz. The values close to these lower bounds were expected only during the supposed /r/ segment. We were careful to view the spectrogram and spectral cross section simultaneously while measuring formant frequencies. We were aware of the possibility of subglottal formants if the voice was breathy and the presence of nasal poles in nasal consonant context. None of the children exhibited hypernasality. The measured formants needed to be continuous between the /r/ and the neighboring vowel. This minimized the possibility of misidentifying a nasal or subglottal formant as a resonance of the supraglottal, non-nasal portion of the vocal tract. The results will be presented in terms of formant frequencies, $F2$ and $F3$, in /r/ and formant separation, $F3-F2$ in /r/. Also, when specific vowel contexts are considered, the differences of $F2$ in /r/ and in the neighboring vowel, and of $F3-F2$ in /r/ and in the neighboring vowel will be discussed.

The first and fourth formant frequencies were not recorded for the children because they simply could not be measured reliably in a sufficient number of tokens. $F1$ was frequently without acoustic energy due to children's high fundamental frequencies: often the fundamental frequency was higher than $F1$; $F4$ was often too faint due to the steep spectral tilt for many of the children's productions. The more reliably measured $F2$ and $F3$ turned out to be sufficiently indicative of differences in /r/ production according to syllable position.

Statistical analyses consisted of linear regression and rank-order analysis. Comparisons of linear regression slopes for these measures across age, grouped according to syllable position, allowed comparisons of the rates of change with age in the formants and formant separations. Because vocal-tract size changes, we can expect that the formant frequencies $F2$ and $F3$ will decrease with age, so that it was best to compare slopes of the regression lines for /r/ in different syllable position. We treated formant separation in a similar way. Because of the variability in the data and the limited number of data points, we accepted an 80% confidence in slope differences as significant. In another kind of analysis an acoustic quantity, such as $F3$, was compared between syllable positions within an age group. In this kind of comparison, rank-order statistics was used because formant frequencies were often distributed non-normally and because they allowed two groups of tokens with relatively low sample sizes to be compared. The probability that two random variables (e.g., $F3$ for /r/ in prevocalic position and $F3$ for /r/ in postvocalic position) possess different probability distributions, even when those distributions are not known, can be tested using rank-order statistics. The particular statistic employed here was the *Wilcoxon* statistic (Bickel and Docksum, 1977, pp. 344–355). In fact, if the hypothesis that the two random variables have the same probability distribution is rejected then one of them is *stochastically greater* than the other. Symbolically, distribution F is stochastically greater than distribution G, if the probability that random X with distribution F is greater than any given t, is greater than the probability of a random variable Y with distribution G being greater than t (Bickel and Docksum, 1977, pp. 344–355; Fisz, 1963, pp. 451–456). In all comparisons using rank-order statistics, the sample sizes were at least ten.

III. RESULTS

The data are presented in categories of age, in months, at recording. Months are grouped by two, except for month 19, when no recording of the eight subjects was made. Using age categories of 2-month duration ensured that enough subjects and recording sessions would be included in each age category. Table I shows the number of samples of prevocalic, postvocalic, medial syllabic, and final syllabic /r/ extracted for analysis for each subject in each age category. Except for one case, this meant that a subject could be included only once in each age category, although not all subjects were included in all age categories. The exception was BT, who was recorded at both 26 and 27 months, and, hence, the number of recording sessions is one more than the number of speakers for this age category.

A. Numbers of words with /r/ or syllabic /r/

Figure 2 shows the mean number of words per recording session, with standard deviations, for analyzed words that had, according to adults' productions, prevocalic /r/, postvocalic /r/, medial syllabic /r/, and final syllabic /r/. These numbers do not indicate whether the children produced perceptually acceptable versions of /r/. The mean number of prevocalic and postvocalic /r/'s per session generally increased across sessions. For postvocalic /r/, the mean incidence increased rapidly over the first three recording sessions, and then remained stable until the last recording session, when there is a dramatic increase again. For prevocalic /r/, the mean incidence remained low until the fourth recording session, when it increased sharply.

There are some important qualifications that should be made about the kinds of tokens that were elicited in the recording sessions. There was no effort made in recording sessions to elicit the same utterances at each session, so there might simply have been variability in the numbers of words with /r/ due to what the parent–child dyad was discussing. For instance, while the rate of final syllabic /r/ word productions increased with age, this was not true for the medial syllabic /r/ production. There was a peak in the latter in the 24–25-month category, with a steady decrease after this age (Fig. 2). A closer look at the data revealed that the frequent producers of medial syllabic /r/'s made many references to the ''Burt'' and ''Ernie'' dolls in the recording room. As a result, it might be the case that rate of medial syllabic /r/ production was actually stable across this period of early language acquisition. Finally, children were dismissed from further recording sessions when they began to routinely produce short (i.e., three-word) sentences. Thus, only the children developing speech at the slowest (normal) rate were included in the last recording session.

B. Children's formant frequencies according to syllable position

Initially, averages of $F2$, $F3$, and $F3-F2$ in /r/ as a function of syllable position and age category were exam-

TABLE I. Table of age categories for which /r/ and syllabic /r/ data were collected for each subject. The numbers in each cell (left to right, top to bottom) represent the number of (1) prevocalic /r/ tokens; (2) postvocalic /r/ tokens; (3) medial syllabic /r/ tokens; and (4) final syllabic /r/ tokens. The bracketed pairs for BT at 26–27 months denote each of these quantities recorded in two separate sessions. "X" is used to indicate that the subject was not recorded for a particular age category.

	15–16	17–18	20–21	22–23	24–25	26–27	28–29	30–31
AN	0, 0, 0, 0	0, 0, 0, 0	3, 6, 1, 0	11, 11, 4, 0	11, 22, 19, 3	X	X	X
BT	0, 1, 0, 0	0, 1, 0, 0	1, 0, 0, 1	X	X	(9,22), (7,7), (8,2), (0,4)	X	X
CK	0, 0, 0, 0	0, 0, 0, 0	0, 0, 0, 0	2, 12, 2, 0	1, 4, 4, 0	X	2, 24, 9, 1	17, 30, 3, 3
DF	0, 0, 0, 0	0, 0, 0, 0	4, 3, 0, 0	9, 7, 3, 1	8, 20, 8, 3	5, 7, 3, 6	21, 1, 3, 6	14, 14, 0, 4
LG	0, 0, 0, 0	0, 0, 0, 0	0, 0, 0, 0	0, 10, 2, 3	9, 16, 4, 7	X	10, 11, 1, 3	21, 20, 7, 19
MS	0, 0, 0, 0	0, 0, 0, 0	2, 3, 1, 0	X	0, 2, 2, 1	6, 8, 2, 0	X	X
MST	0, 0, 0, 0	3, 0, 0, 0	9, 4 16, 3	X	15, 14, 20, 9	11, 16, 17, 5	13, 11, 7, 10	X
RF	0, 0, 0, 0	0, 0, 0, 0	2, 3, 0, 0	2, 3, 6, 2	0, 0, 3, 0	25, 15, 1, 5	22, 8, 0, 12	8, 9, 0, 2
Total number of subjects	8	8	8	5	7	6	5	4

ined. Thus, the data from different speakers and different vowel environments were pooled for the analyses in this section. For both prevocalic and postvocalic /r/ [Figs. 3(a) and (b)] there was a slight tendency for $F2$ and $F3$ to decrease with age. For $F2$ the trend appeared to be stronger for prevocalic /r/ than for postvocalic /r/: Between 26 and 31 months the mean $F2$'s for prevocalic /r/ was between 1600 and 1900 Hz, while for postvocalic /r/ the mean $F2$ was between 1900 and 2400 Hz. For $F3$ the trend appeared stronger for postvocalic /r/ than it was for prevocalic /r/: Between 26 and 31 months the mean $F3$ for postvocalic /r/ was about 3200 Hz, while it was about 3700 Hz for prevocalic /r/. The result of these trends was a faster decrease in mean $F3-F2$ for postvocalic /r/ than for prevocalic /r/ with age category. There did not appear to be a consistent change with age category in mean $F3-F2$ for prevocalic /r/: This quantity stayed close to 2000 Hz throughout the age categories except for 14–15 months, where it was even higher. On the other hand, there was a general downward trend in this quantity with age for postvocalic /r/, so that by the 28–29-month and 30–31-month categories mean $F3-F2$ was closer to 1000 Hz. A 1000-Hz difference in frequencies is less than would be expected for a neutral vowel for ages 1 to 2 1/2 years.[2]

Linear regression was performed on $F2$, $F3$, and

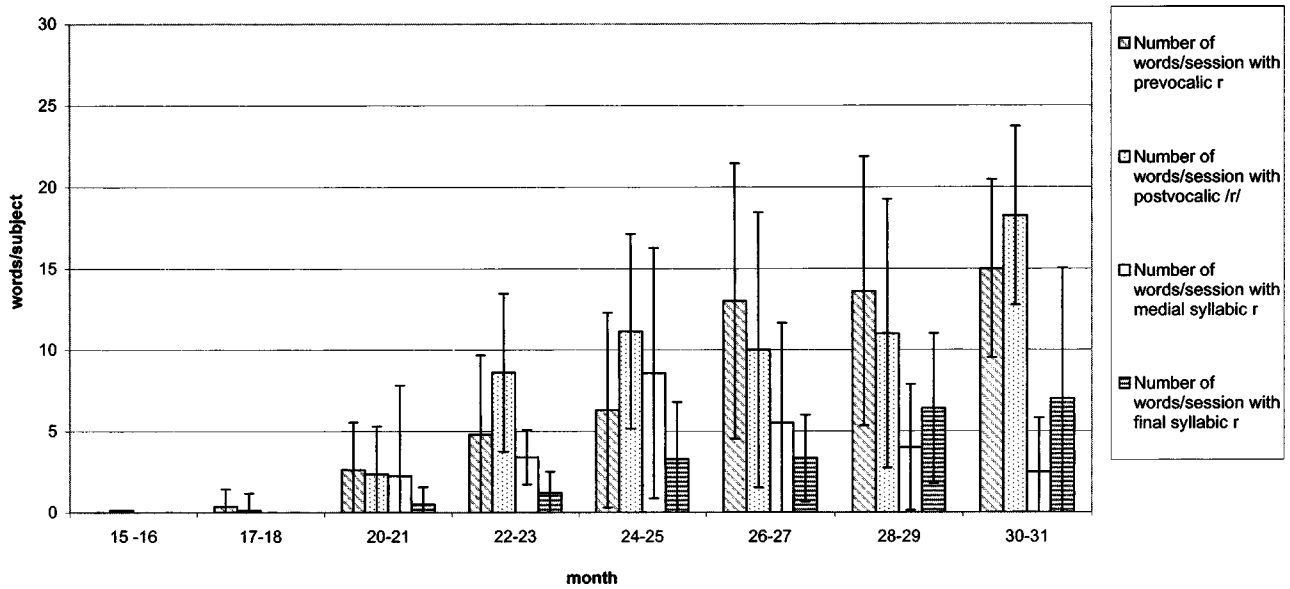


FIG. 2. Mean numbers and standard deviations of analyzed words per recording session as a function of age category. Words either possess prevocalic /r/, postvocalic /r/, medial syllabic /r/, or final syllabic /r/.

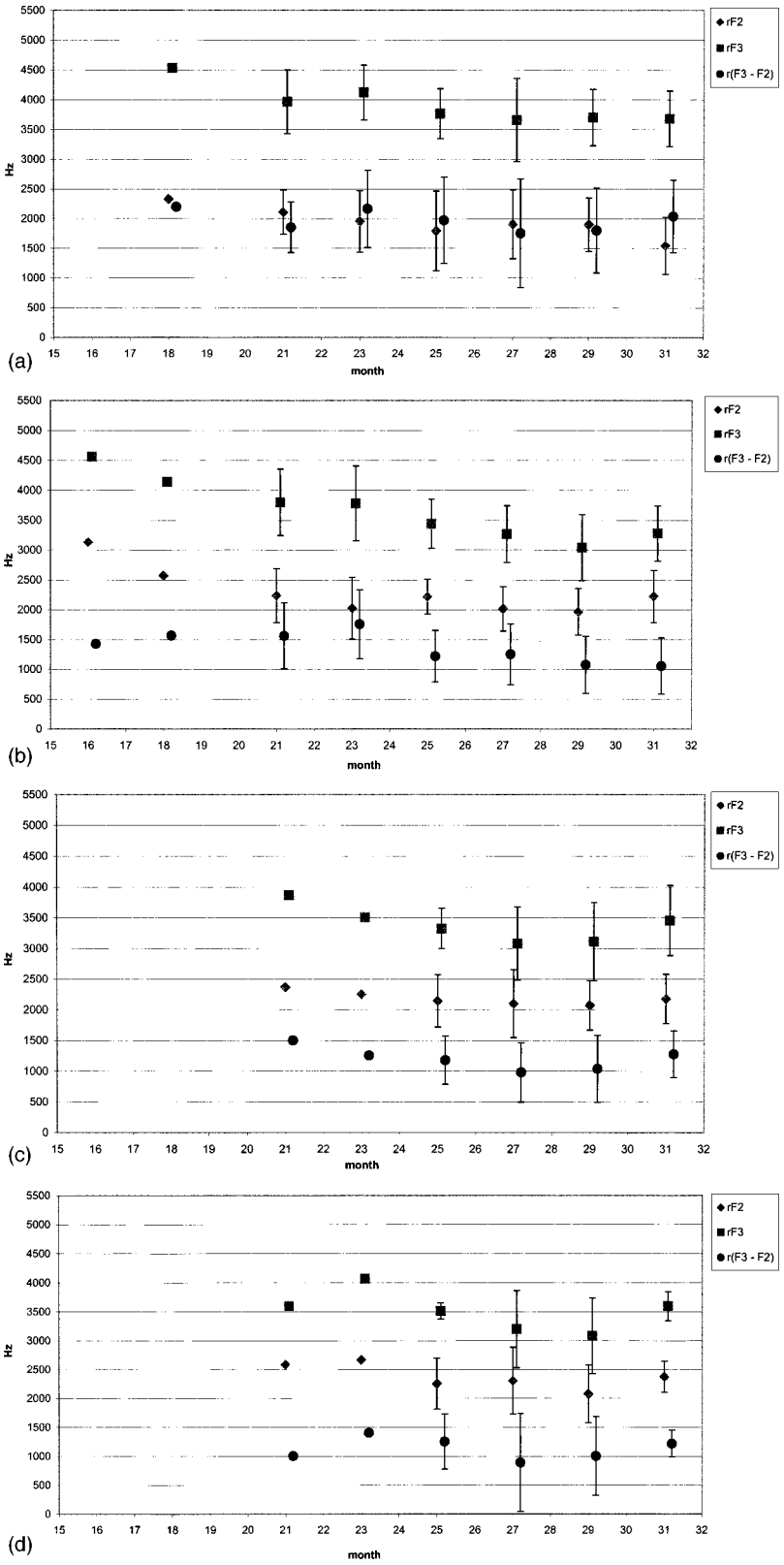


FIG. 3. (a) Means and standard deviations of $F2$, $F3$, and $F3-F2$ in prevocalic /r/ for the children subjects as a function of age category. (b) Means and standard deviations of $F2$, $F3$, and $F3-F2$ in postvocalic /r/ for the children subjects as a function of age category. (c) Means and standard deviations of $F2$, $F3$, and $F3-F2$ in medial syllabic /r/ for the children subjects as a function of age category. (d) Means and standard deviations of $F2$, $F3$, and $F3-F2$ in final syllabic /r/ for the children subjects as a function of age category.

$F3-F2$ for prevocalic, postvocalic, medial syllabic, and final syllabic /r/ versus age, in months. The slopes of the regression lines and their 80% confidence intervals are tallied in Table II. This analysis revealed that the 80% confidence ranges for the slope of the postvocalic $F3-F2$ versus age line and the prevocalic $F3-F2$ versus age line did not overlap (Table II, rows 1 and 2, column 3): The rate of decrease in $F3-F2$ was greater for postvocalic /r/ than for prevocalic /r/. Most of the difference between syllable position for the rate of change of $F3-F2$ appears to be the more rapid decline of $F2$ with age in prevocalic position than in postvocalic position, and the more rapid decline of $F3$ with age in postvocalic position than in prevocalic position (Table II, rows 1 and 2, columns 1 and 2). There was, however, overlap

TABLE II. Slopes and their 80% confidence intervals for linear regression of formant frequencies (Hz) in /r/ versus age, in months, of recording for /r/ in various syllable positions.

	$F2$	$F3$	$(F3-F2)$
Prevocalic	−35.0	−36.1	−1.1
	[−47.3,−22.6]	[−48.6,−23.6]	[−18.2,15.9]
Postvocalic	−14.7	−48.0	−33.3
	[−23.6,−5.8]	[−59.0,−37.0]	[−44.2,−22.3]
Medial syllabic	−34.2	−53.2	−19.0
	[−47.6,−17.0]	[−78.1,−28.4]	[−41.0,3.0]
Final syllabic	28.7	−15.1	−43.8
	[11.1,46.1]	[−42.8,12.6]	[−72.4,−15.2]

in the 80% confidence intervals for the prevocalic and postvocalic slopes for both $F2$ and $F3$. None of the slope comparisons showed differences at the 90% confidence level.

Figure 3(c) shows $F2$, $F3$, and $F3-F2$ as a function of age for medial syllabic /r/. Except for the 30–31-month category, there was an overall downward trend from 20–21 months in $F2$, $F3$, and $F3-F2$ for medial syllabic /r/ as a function of age category. Medial syllabic /r/ behaved similarly to postvocalic /r/ with mean $F3-F2$ at about 1500 Hz in the 20–21-month category and near 1000 Hz in the 26–27- and 28–29-month categories. Similarly, the mean $F3$ fell from about 3800 Hz at 20–21 months to 3200 Hz at 26–27 months. Table II shows that the 80% confidence intervals for the rate of decline of the formant frequencies with age for medial syllabic /r/ overlapped the 80% confidence intervals for these rates for both the prevocalic and postvocalic /r/'s (Table II, rows 1, 2, and 3).

For final syllabic /r/ [Fig. 3(d)], the mean $F3$ was between 3000 and 3600 Hz from 24–25 months through 30–31 months, without an apparent trend with age. For the same span of time, $F3-F2$ was between 800 and 1300 Hz. The regression analysis showed that $F2$, in fact, increased with age, while $F3-F2$ decreased with age (Table II, row 4). The 80% confidence intervals for the positive slope of $F2$ versus age for final syllabic /r/ did not overlap with any of the 80% confidence intervals for the other syllable positions. While the negative slope for $F3-F2$ in final syllabic /r/ possessed the largest absolute value of the four types of /r/, the 80% confidence interval was quite broad.

Wilcoxon statistics were employed to characterize the differences in the distributions of the formant frequencies between types of /r/ in the 30–31-month age category. Table III shows the probability: (1) that the $F2$'s for postvocalic, medial vocalic, and final vocalic /r/'s possessed distributions stochastically greater than that for $F2$'s in prevocalic /r/'s, and (2) that the $F3$'s and $(F3-F2)$'s for postvocalic, medial vocalic, and final vocalic /r/'s possessed distributions stochastically less than those for prevocalic /r/'s in the 30–31-month age category. The $F2$'s for postvocalic, medial syllabic, and final syllabic /r/ all possessed distributions stochastically greater than that for prevocalic $F2$ with a very high probability (Table III, column 1). Further, the distributions of $F3$'s and $(F3-F2)$'s for postvocalic and final syllabic /r/'s were stochastically less than the corresponding distributions for prevocalic /r/ with a very high probability (Table III, columns 2 and 3, rows 1 and 3). On the other hand, there was almost no certainty that the distribution of $F3$ for medial syllabic /r/ was stochastically less than the distribution of $F3$ for prevocalic /r/ (Table III, column 2, row 2). This appeared to have had some effect on the probability that the distribution of $F3-F2$ for medial syllabic /r/ was stochastically less than that for prevocalic /r/ (Table III, column 3, row 2). However, it should be kept in mind that the number of tokens of medial syllabic /r/ was dwindling by the 30–31-month age category (Fig. 2).

TABLE III. The probability, to the nearest 0.001: (1) that the $F2$'s for postvocalic, medial vocalic, and final vocalic /r/'s possess distributions stochastically greater than that for prevocalic /r/'s, and (2) that the $F3$'s and $(F3-F2)$'s for postvocalic, medial vocalic, and final vocalic /r/'s possess distributions stochastically less than those for prevocalic /r/'s in the 30–31-month age category.

Prevocalic versus	$F2$	$F3$	$(F3-F2)$
Postvocalic	>0.999	>0.999	>0.999
Medial syllabic	0.999	0.599	0.985
Final syllabic	>0.999	>0.999	>0.999

Some of the prevocalic /r/'s appeared within consonant clusters. Hoffman, Schuckers, and Ratusnik (1977) found that certain initial stop consonants could facilitate the correct production of prevocalic and vocalic /r/ in children from about 6 years to 7 years of age. To test the effect that the alveolar and velar stop consonants had on the formant frequencies in /r/, rank-order analysis with the Wilcoxon statistic was performed comparing prevocalic /r/'s in consonant clusters with alveolar and velar stops and singleton prevocalic /r/'s in the age range from 28 to 31, months. Prevocalic /r/'s in consonant clusters with just alveolar stops and singleton prevocalic /r/'s in the age range from 28 to 31 months were also compared. In both cases, the prevocalic /r/'s in consonant clusters had distributions with $F3-F2$ stochastically less than the distribution for singleton prevocalic /r/'s, with small to moderate probabilities (0.81 for clusters with alveolar and velar stops versus singletons and 0.66 for clusters with alveolar stops versus singleton). However, the prevocalic /r/'s in consonant clusters had distributions with $F3$ stochastically greater than the distributions for singleton prevocalic /r/'s, with moderate probability (0.90 for clusters with alveolar and velar stops versus singletons and 0.91 for clusters with alveolar stops versus singleton). In these data, alveolar and velar stop consonants did not appear to promote correct prevocalic /r/ production.

C. Children's formant frequencies according to vowel context and syllable position

A more detailed analysis of the nature of prevocalic and postvocalic /r/ production can be attained when the word tokens are segregated according to both syllable position and according to the identity of the vowel neighboring the /r/. The neighboring vowel succeeds a prevocalic /r/ and precedes a postvocalic /r/. When neighboring vowels are considered, differences between the formant frequencies in /r/ and in the neighboring vowel can be calculated. This allows a consideration of the formant trajectories between /r/ and its neighboring vowel, which further characterizes the prevo-

TABLE IV. Number of tokens of prevocalic and postvocalic /r/ in vowel context as a function of age category.

Months	Prevocalic /r/ in /ɛ/ context	Prevocalic /r/ in /ɑ/ context	Postvocalic /r/ in /ɛ/ context	Postvocalic /r/ in /ɑ/ context
15–16	0	0	1	0
17–18	0	1	1	0
20–21	1	2	5	4
22–23	1	7	25	11
24–25	4	2	39	19
26–27	5	4	17	16
28–29	4	6	18	18
30–31	8	0	40	6

calic and postvocalic /r/. For example, the differences in $F2$ in /r/ and the $F2$ in the neighboring vowel can be considered. Also, differences in formant frequency separation, i.e., differences in $F3-F2$, for /r/ and for the neighboring vowel can be considered. First, productions of /r/ in prevocalic and postvocalic position in two different vowel contexts are discussed. Then, briefly, the formant frequency changes during syllabic /r/ are examined.

The vowels chosen for examination were the most common in the subjects' productions with prevocalic and postvocalic /r/. These vowels were /ɛ/ and /ɑ/ as in the words ''red'' and ''frog,'' respectively, or the words ''there'' and ''car,'' respectively. Table IV shows the number of tokens of each type of /r/ as a function of age category. Focusing on age categories greater than 20–21 months, Fig. 4 reveals that the mean separation between the third and second formant frequencies, $F3-F2$, was actually larger in the prevocalic /r/ than in a succeeding /ɛ/, while this separation was constant in the mean between a postvocalic /r/ and its preceding /ɛ/. Formant separation behavior was different for the back vowel /ɑ/ compared to that of the front vowel /ɛ/. $F3-F2$ was constant in the mean from a prevocalic /r/ to its succeeding /ɑ/, but there was a decrease in mean formant separation (i.e., smaller $F3-F2$) from an /ɑ/ to its succeeding /r/ (Fig. 4). Also, the behavior of $F2$ depended on whether the neighboring vowel was the front or the back vowel. Figure 5 shows that mean $F2$'s tended to be lower in both the prevocalic and postvocalic /r/ than in the neighboring front vowel /ɛ/. For the back vowel /ɑ/ the mean $F2$'s for prevocalic /r/ were lower than the succeeding /ɑ/, while mean $F2$'s in postvocalic /r/ were higher than in the preceding /ɑ/. These observations of $F2$ help to explain the differences in formant separation, $F3-F2$ behavior in front vowel /ɛ/ context and back vowel /ɑ/ context.

Wilcoxon statistics were used to compare prevocalic /r/ to postvocalic /r/ in /ɛ/ and /ɑ/ contexts. For the /ɛ/ context tokens were from the ages of 28 through 31 months, and for /ɑ/ the tokens were from the ages of 26 through 29 months. (These age categories were chosen to include a sufficient number of tokens to perform statistics, at the same time attempting to limit the range of variation caused by vocal-tract length changes.) Column 1 of Table V shows that the distributions for $F2$ in postvocalic /r/ were stochastically greater than those for prevocalic /r/ with a high probability in both /ɛ/ and /ɑ/ contexts. Similarly, the distributions for $F3$ and $F3-F2$ in postvocalic /r/ were stochastically less than for prevocalic /r/, with a high probability in both vowel contexts (Table V, columns 2 and 3). It can be noted that the probabilities regarding the $F2$ distributions are greater for /ɛ/ than for /ɑ/, and those pertaining to the $F3$ distributions are greater for /ɑ/ than for /ɛ/. This seems to have balanced out to produce very high probabilities for differences between postvocalic and prevocalic /r/ in the distributions for $F3-F2$. Columns 4 and 5 of Table V show that the distributions for the differences between $F3$ and $F3-F2$ in the /r/ and in the vowel were stochastically less than, or stochastically more negative, for postvocalic /r/ than for prevocalic /r/. That is, the *absolute* differences between $F3$ and $F3-F2$

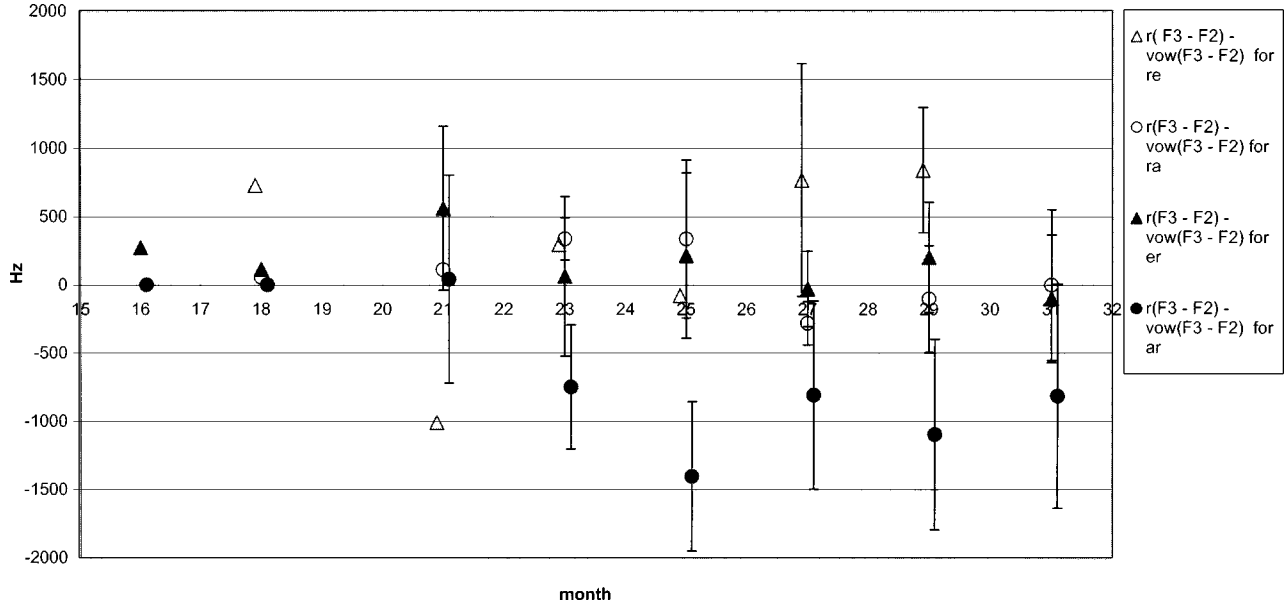


FIG. 4. Means and standard deviations of differences of $F3-F2$ values in the /r/ and the neighboring vowel for prevocalic and postvocalic /r/ in /ɛ/ and /ɑ/ context for the children subjects as a function of age category.

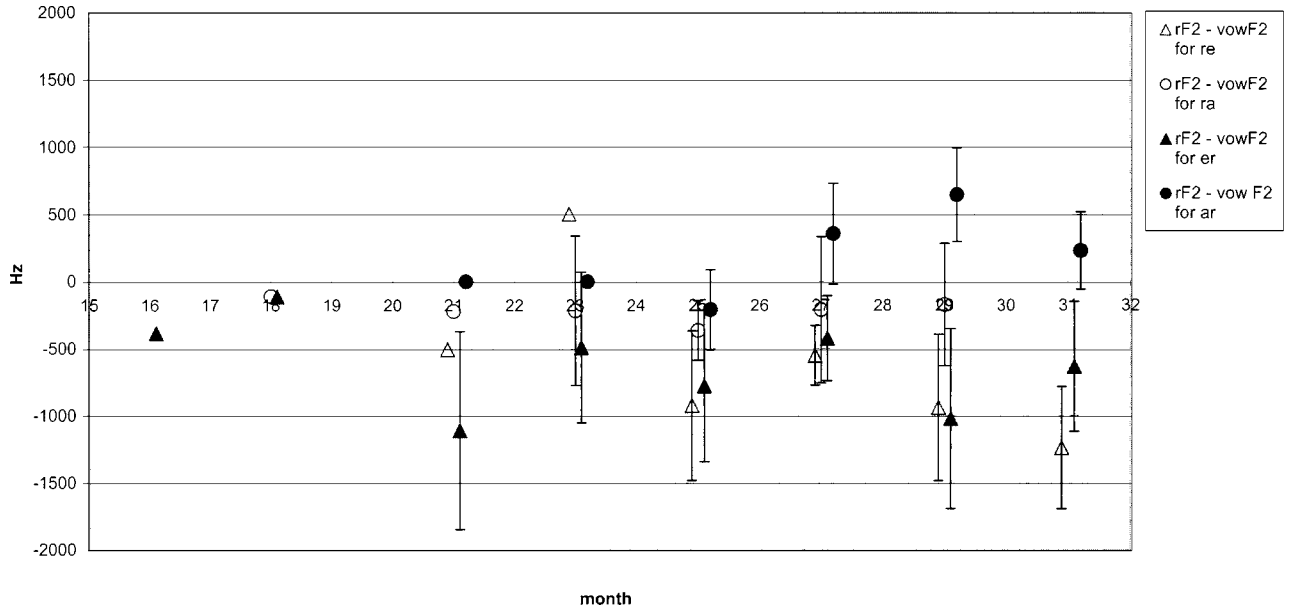


FIG. 5. Means and standard deviations of differences of $F2$ values in the /r/ and the neighboring vowel for prevocalic and postvocalic /r/ in /ɛ/ and /ɑ/ context for the children subjects as a function of age category.

in /r/ and in the vowel tended to be *greater* for postvocalic /r/ than for prevocalic /r/.

The syllabic /r/'s were produced as monophthongs, almost always. Changes in $F2$, $F3$, and $F3-F2$ from the beginning of syllabic /r/ and to the end of each medial and final syllabic /r/ were computed. The mean change in these formant measures was never greater than 500 Hz in magnitude, and most often less than 300 Hz. Further, zero frequency change was within 1 standard deviation of the mean, except in three tokens: (1) one token of $F3-F2$ at 20–21 months for medial syllabic /r/; (2) one token of $F2$ and $F3-F2$ at 14–15 months for final syllabic /r/, and (3) one token of $F3$ at 28–29 months for final syllabic /r/. In the first and third cases, zero frequency change was well within 2 standard deviations of the mean.

IV. DISCUSSION

The results indicate that this group of children was progressing toward postvocalic [ɹ] more rapidly than they were progressing toward prevocalic [ɹ]. The children's $F3-F2$ declined with age faster for postvocalic /r/ than for prevocalic /r/. This result appears to be due to a decrease of $F3$ and a more gradual decrease of $F2$ for postvocalic /r/ with age in comparison with age-related changes in these formant frequencies for prevocalic /r/ (Table II). Wilcoxon statistics showed that the distributions for $F3-F2$ for final syllabic /r/ and for postvocalic /r/ at 30–31 months were stochastically smaller than the corresponding distributions for prevocalic /r/ with very high probabilities (>0.999) (Table III). While $F3-F2$ of final syllabic /r/ appeared to decrease rapidly with age, there was great variability associated with the slope of the regression line (Table II). On the other hand, $F2$ for final syllabic /r/ increased with age, providing for some of the decrease in $F3-F2$. The behavior of the formant measures for medial syllabic /r/ was more equivocal, probably because the number of tokens of these became relatively small after the 26–27-month category. The functional criterion for including children in the study group (Sec. II A) may have had some effect on the trends in time. However, the comparisons between groups at any age are valid for the particular children at that time.

In a comparison of front and back vowel contexts, /ɛ/ and /ɑ/, it was seen that vowels did not change the statistical relations in the formant frequencies between prevocalic and postvocalic /r/ (Table V). There were, however, effects of vowel context on the $F2$ and formant frequency separations, $F3-F2$ trajectories. For instance, Fig. 5 shows that $F3-F2$ tended to decrease from /ɑ/ into postvocalic /r/, but not for

TABLE V. The probability, to the nearest 0.001: (1) that the $F2$'s for postvocalic /r/'s possess distributions stochastically greater than that for prevocalic /r/'s, and (2) that the $F3$'s and $(F3-F2)$'s for postvocalic /r/ and the differences with their values in the neighboring vowel possess distributions stochastically less than those for prevocalic /r/'s. The age range is 28 to 31 months in the context of vowel /ɛ/ and it is 26 to 29 months in the context of vowel /ɑ/.

Prevocalic versus	$F2$ in /r/	$F3$ in /r/	$(F3-F2)$ in /r/	$F3$ in /r/- $F3$ in vowel	$(F3-F2)$ in /r/- $(F3-F2)$ in vowel
Postvocalic in /ɛ/ context	0.999	0.995	>0.999	0.994	>0.999
Postvocalic in /ɑ/ context	0.991	>0.999	>0.999	0.996	0.999

/ɛ/ into postvocalic /r/. A rising $F2$ from /ɑ/ into postvocalic /r/ could contribute to a decreasing formant separation, $F3-F2$. This is consistent with adult formant trajectories for final [ɹ] (e.g., Olive *et al.*, 1993, p. 222), and would indicate a tongue becoming less backed in the transition from the vowel into the liquid. Also, the children seemed to have been producing medial and final syllabic /r/'s as monophthongs, as would be expected in adult production.

The children for whom data are reported were, on the whole, progressing toward postvocalic [ɹ] and final syllabic [ɹ] more rapidly than they were progressing toward prevocalic [ɹ]. While there was acoustic evidence that progress was made by some subjects toward a prevocalic [ɹ], the group as a whole showed little evidence of this. The question naturally arises whether the differences in the acquisition of postvocalic and syllabic versus prevocalic [ɹ] are the result of perceptual or production mechanisms. Further, precisely where is the difficulty? A definite answer to this question cannot be given, but there are some useful pieces of evidence that can be used to argue for certain causes.

In the review of the literature on adult [ɹ] production and in the case of the adult speaker from the XRMB-SPD, prevocalic [ɹ] was identified as having lower $F1$, $F2$, and $F3$ than postvocalic [ɹ], and the prevocalic [ɹ] appears to require two substantial tongue constrictions: one in the oral cavity and the other in the pharynx. The postvocalic [ɹ] and syllabic [ɹ] apparently do not require a tight, or any, pharyngeal constriction. Based on the fact that the newborn infant's tongue fills the oral cavity (Kent and Vorperian, 1995), we would expect that the ratio of the tongue volume to the volume of the supralaryngeal vocal tract is larger for infants and young children than it is for adults. This, together with the fact that the larynx is descending rapidly in the age group considered here (Goldstein, 1980; Kent and Vorperian, 1995), could mean that articulations appropriate for any prevocalic [ɹ] are difficult to attain. Both a bulky tongue body and a small pharyngeal cavity would hinder young children's ability to form both a palatal and pharyngeal constriction with the tongue.

An argument could also be made that the motor control is not mature enough for prevocalic [ɹ] production before, say, 2 or 2 1/2 years. One aspect is that the control of the tongue blade requires some time to mature. Children who are native Spanish speakers master the trilled /r/ at a relatively late age (Jimenez, 1987). We have also noticed that children aged 3 to 7 years make less of an acoustic distinction between /s/ and /ʃ/ than do adults (McGowan and Nittrouer, 1988; Nittrouer, 1995), despite the fact that they can perceive the two phonemes categorically. However, these differences could also be due to the morphology in this region of the tongue and palate. For instance, young children may not be able to create a sufficiently large sublingual cavity for a retroflex [ɹ], just as they have difficulty in producing one for [ʃ] (Perkell, Boyce, and Stevens, 1979; Nittrouer, 1995).

The role of perception in this story is not known, yet its influence cannot be discounted even when there is a plausible articulatory explanation for an observed speech acoustics trend. In fact, there has been some research on the perception and production link of [ɹ] in young children. Menyuk and Anderson (1969) found that their preschool subjects had category boundaries in forced-choice identification between prevocalic /r/, /l/, and /w/. However, according to adult judges, when the children were asked to repeat the words they perceived to be ''write'' they very often reproduced the word with something close to ''white.'' In general, there was a mismatch between the perceived phoneme and the reproduced phoneme, at least to adult ears, particularly for the /w/–/r/ continuum. The results of the Menyuk and Anderson work indicate some decoupling between perception and production of word initial /r/ in young children. On the other hand, Strange and Broen (1981) found that many of their 3-year-old subjects who tended to produce the word-initial /r/ poorly also tended to be less adult-like in categorization tasks. Thus, the Strange and Broen work emphasized a certain amount of coupling between production and perception capability for word-initial /r/ in children. It should be noted that Strange and Broen used more sophisticated stimuli in simulating their /w/–/r/ continua than did Menyuk and Anderson. The evidence indicates that there is some causal connection between perceptual and production capabilities in learning to produce prevocalic [ɹ], but that one capability does not determine the other. Even for adults who are physiologically capable of producing prevocalic [ɹ] may be able to hear postvocalic [ɹ] more easily than they can hear prevocalic [ɹ], thus leaving them unable to produce prevocalic [ɹ]. This could account for the differences in Japanese identification of [ɹ] and [l] based on syllable position found by Mochizuki (1981).

The previous data of others and current data presented here lead to interesting directions for future research. Below, we briefly discuss some of the speculations and questions raised here. In contrast to prevocalic [ɹ], the young children in this data set were achieving low $F3$ and $F3-F2$ appropriate for postvocalic [ɹ]. Also, in general, children do not have difficulty producing [w] (e.g., Smits, 1993). Three low formants are required for prevocalic [ɹ], while postvocalic [ɹ] and [w] possess only one or two particularly low formants: $F1$ and $F2$ for [w], and $F3$ for postvocalic [ɹ]. Further, postvocalic [ɹ] and [w] can be articulated using only one oral constriction, other than the lips. The [w] can be articulated using one tongue-body approximation in the velar region and lip rounding. For [w], the first two formants can be modeled using two coupled Helmholtz resonators, and the third formant is the half-wave resonance of either the front or back cavities (Stevens, 1998). Young children should also be able to produce the simple two-cavity system, even if it entails a different cavity affiliation for the third formant than for adults. Similarly, we suggest that postvocalic [ɹ], and syllabic [ɹ] are produced with a single tongue constriction, which is more forward than that for [w]. Particularly, postvocalic [ɹ] behaves as an off-glide to the preceding vowel that young children are able and willing to produce.

Why, if young children can produce postvocalic /r/ with relatively low $F3$ and small $F3-F2$, don't they do so for prevocalic /r/? In fact many children substitute [w] for prevocalic [ɹ] (Shriberg and Kent, 1995). We speculate that young children tend to substitute [w] for [ɹ] in prevocalic position because they can only produce two low-formant fre-

TABLE VI. The ranges of formant values measured in [ɹ] or syllabic [ɹ] for tokens of several words for JW11 in the XRMB-SPD. The numbers in parentheses are the numbers of tokens analyzed.

	$F1$ range (Hz)	$F2$ range (Hz)	$F3$ range (Hz)	$F4$ range (Hz)	Articulatory properties
"Right" (3)	331 to 441	828 to 961	1255 to 1435	2262 to 2373	Retroflex, lip rounding
"Rag" (2)	386 to 497	938 to 1048	1545 to 1655	2759 to 2814	Nonretroflex, lip rounding tongue distance-to-palate <0.6 cm
"There" (1)	607	1545	1931	2869	Nonretroflex, no lip rounding tongue distance-to-plate >0.8 cm
"Large" (3)	534 to 694	1324 to 1389	1710 to 1843	2814 to 2924	Nonretroflex, no lip rounding tongue distance-to-palate >0.8 cm
"Dormer" (1st syllable) (2)	607 to 662	1102 to 1159	1766 to 1876	2759	Nonretroflex, no lip rounding tongue distance-to-palate >0.8 cm
"Dormer" (2nd syllable) (2)	607	1159 to 1214	1655	2704	Nonretroflex, no lip rounding tongue distance-to-palate <0.4 cm

quencies simultaneously. A reasonable compromise would be achieved in this substitution, because prevocalic [ɹ] possesses relatively low $F1$ and $F2$, and these formants are, perhaps, perceptually more salient than $F3$. Some children appear to make this compromise in articulatory behavior and continue it even after they are physically more capable of producing a prevocalic [ɹ]. The reason that prevocalic [ɹ] is sometimes not forthcoming, even when their productive lexicons requires an /r/–/w/ distinction, is a topic for further research. Perhaps part of the answer is that some of these children make a subphonemic distinction between prevocalic /w/ and /r/, which is an acceptable categorical distinction for them (Hoffman, Stager, and Daniloff, 1983). Also, children who do not start toward an adult-like prevocalic [ɹ] at an early enough age may have difficulty in changing a strong coupling between learned motor behavior and perceptual attention.

V. CONCLUSION

The formant frequency data on this group of children, from about 14 months through 26 months, and some through 31 months, indicated that they were acquiring aspects of [ɹ] production for postvocalic and syllabic /r/ before they were acquiring equivalent aspects of production for prevocalic /r/. That is, the distributions of $F3$'s and $(F3-F2)$'s for postvocalic, and least for final syllabic, /r/ were stochastically less than those for prevocalic /r/ in the final months of recording. Also, the rate of decline for $F3-F2$ with age was significantly greater in postvocalic /r/ than in prevocalic /r/. However, the prevocalic /r/ $F2$ distributions were stochastically smaller than those for postvocalic and syllabic /r/'s with a very high probability, which is also typical of the differences seen in adults according to syllable position. A comparison of a front and a back vowel context did not reveal that vowel context affected these basic results. We have attributed part of the reason for lack of progress in prevocalic /r/ production to the complexity of its articulation. However, this is not the complete explanation and more research into the relation between perception and production needs to be pursued for a complete picture.

ACKNOWLEDGMENTS

This work was supported by research Grant No. R01 DC 00633 from the National Institute on Deafness and Other Communication Disorders to the second author. We thank Karen Chenausky for comments on an earlier version of this paper, and Melanie Wilhelmsen for help with editing. The comments of two anonymous reviewers helped to improve this paper.

APPENDIX: A CASE STUDY ON AN ADULT

The subject JW11 from the XRMB-SPD was examined in some of his productions of [ɹ], because he used both retroflex and nonretroflex articulations to produce prevocalic [ɹ]. Table VI presents the ranges of the first four formant frequencies and their articulatory correlates for JW11 in [ɹ] for each of the words "right," "rag," "there," "large," and "dormer."

There was a tendency for all formant frequencies to be lower for the retroflex prevocalic [ɹ] than for nonretroflex prevocalic [ɹ] for all formants, as shown in rows 1 and 2 of Table VI. This was particularly clear for $F4$, and this corroborates the 1999 finding of Espy-Wilson and Boyce.

Each postvocalic [ɹ] was produced with a nonretroflex articulation, and formant frequencies, are shown in rows 3 through 5 of Table VI. The first three formant frequencies for all postvocalic [ɹ]'s were higher than the corresponding formant frequencies for both the retroflex and nonretroflex articulations of prevocalic [ɹ]. In particular, even with the relatively low $F2$, caused by the backing of the tongue for the preceding vowel in "dormer," the $F2$'s for postvocalic [ɹ]'s were all greater than for the prevocalic [ɹ]'s. The $F4$'s for postvocalic [ɹ] were similar to those of the nonretroflex prevocalic [ɹ].

The second syllable of the two tokens of "dormer" were examined as examples of syllabic [ɹ] in row 6 of Table VI. These were produced as nonretroflex articulations and only slightly diphthongized. Except for the $F3$'s as low as the highest value for a prevocalic [ɹ] production, the formant frequency ranges were in similar relation to those of the prevocalic [ɹ]'s as were those of the postvocalic [ɹ]'s.

Pellet positions were examined to discover the articulatory bases for the measured formant frequencies. JW11 produced lip rounding during prevocalic [ɹ], but not for postvocalic or syllabic [ɹ]. This helped to account for at least some of the differences in formant frequencies between nonretroflexed prevocalic [ɹ] and postvocalic [ɹ].

In consideration of articulatory factors, the nonretroflexed [ɹ] is examined first. The two chosen examples of the word ''rag'' spoken by JW11 possessed nonretroflex [ɹ]. The constriction for [ɹ] was forward of the constriction for [g] (i.e., it was more palatal than velar). However, the tongue was more ''bunched'' for [ɹ] compared to [g], so that the tongue tip was further back and higher for [ɹ] than for [g]. Also, lip rounding occurred for prevocalic [ɹ], but not for [g]. This caused the $F3$ for [ɹ] to be lower than for [g] because the front cavity served as a Helmholtz resonator with resonance frequency $F3$ in [ɹ] whose volume (capacitance) and mass elements are maximized. (The $F3$ [g] was measured during the burst.) On the other hand, for [g] in a front vowel context, $F3$ can be approximately a quarter-wave resonance of the front cavity. The predictions for $F2$ in these segments are more difficult because there is no indication of the tongue root or larynx position in the XRMB-SPD data set. In fact, $F2$ for [ɹ] was substantially lower (938 and 1048 Hz) than it was for [g] (1644 and 1821 Hz), which is consistent with the palatal constriction being more forward and the tongue more bunched for [ɹ] than for [g].

What was the difference in JW11's articulation between nonretroflex prevocalic [ɹ], and postvocalic and syllabic [ɹ], other than lip rounding? An examination of the articulatory data reveals that the difference was partly one of palatal constriction degree, as measured by the minimum distance of the pellets to the palate. The palatal approximation that the tongue made was not as narrow for the postvocalic [ɹ] in ''there,'' ''large,'' and ''dormer'' as it is for the nonretroflexed prevocalic [ɹ] in ''rag.'' In the midsagittal plane, postvocalic [ɹ] had a minimum postalveolar pellet-to-palate distance of at least 0.8 cm, while for prevocalic [ɹ] this distance was less than 0.6 cm. A lower tongue position for postvocalic [ɹ] compared to nonretroflex prevocalic [ɹ] is consistent with the higher $F1$ in postvocalic [ɹ] observed for JW11. The higher $F3$ in postvocalic [ɹ] compared to prevocalic [ɹ] is consistent with reduced lip rounding and palatal constriction degree in the former compared to the latter. Perhaps because the syllabic [ɹ] tokens from ''dormer'' tokens were the final syllable in a word from a word list, these were more strongly articulated than they would be in other situations. The $F3$'s from the syllabic [ɹ]'s were low compared to those of the postvocalic examples from JW11. Views of JW11's articulations reveal that these syllabic [ɹ]'s are produced with tight postalveolar or palatal constrictions (pellet-to-palate distances were less than 0.4 cm). The moderate $F1$'s may have been the result of a reduced palatal constriction length that decreased the effective mass of the Helmholtz resonator that gave rise to $F1$.

[1]The various phonetic variants of American English /r/ will be denoted collectively by the symbol [ɹ]. The phoneme symbol /r/ will be employed when discussing children's productions of words that would contain [ɹ] in an adult's production, rather than referring to ''attempted productions of [ɹ].''

[2]The spacing between adjacent formants should be at least 1400 Hz for 4-year-old children, based on data provided by Kent and Forner (1979). This estimate is based on a formant scale factor of at least 40% and an average adult male frequency spacing of 1000 Hz. The formant spacing should be even greater for children less than 2.5 years old. The same data indicate that the $F3$ for neutral vowels is greater than 3500 Hz for the group of children considered here.

Alwan, A., Narayanan, S., and Haker, K. (**1997**). ''Toward articulatory–acoustic models for liquid approximants based on MRI and EPG data. II. The rhotics,'' J. Acoust. Soc. Am. **101**, 1078–1089.

Bickel, P. J., and Doksum, K. A. (**1977**). *Mathematical Statistics: Basic Ideas and Selected Topics* (Holden-Day, San Francisco).

Dalston, R. (**1975**). ''Acoustic characteristics of English /w,r,l/ spoken correctly by young children and adults,'' J. Acoust. Soc. Am. **57**, 462–469.

Delattre, P., and Freeman, D. C. (**1968**). ''A dialect study of American r's by X-ray motion picture,'' Linguistics **44**, 29–68.

Espy-Wilson, C. Y., and Boyce, S. (**1999**). ''The relevance of $F4$ in distinguishing between different articulatory configurations of American English,'' J. Acoust. Soc. Am. **105**, 1400.

Espy-Wilson, C. Y., Boyce, S. E., Jackson, M., Narayanan, S., and Alwan, A. (**2000**). ''Acoustic modeling of American English /r/,'' J. Acoust. Soc. Am. **108**, 343–356.

Fisz, M. (**1963**). *Probability Theory and Mathematical Statistics*, 3rd ed. (Wiley, New York).

Goldstein, U. G. (**1980**). ''An articulatory model for the vocal tracts of growing children,'' Ph.D. dissertation, MIT, Cambridge, MA).

Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, M., and Perkell, J. S. (**1999**). ''Articulatory tradeoffs reduce acoustic variability during American English /r/ production,'' J. Acoust. Soc. Am. **105**, 2854–2865.

Hagiwara, R. (**1994**). ''Three types of American /r/,'' UCLA Working Papers of Phonetics, **88**, 55–61.

Hoffman, P. R., Schuckers, G. H., and Ratusnik, D. L. (**1977**). ''Contextual-coarticulatory inconsistency of /r/ misarticulation,'' J. Speech Hear. Res. **20**, 631–643.

Hoffman, P. R., Stager, S., and Daniloff, R. G. (**1983**). ''Perception and production of misarticulated /r/,'' J. Speech Hear Disord. **48**, 210–248.

Jackson, M. T.-T., Espy-Wilson, C. Y., and Boyce, S. E. (**2001**). ''Verifying a vocal tract model with a closed side-branch,'' J. Acoust. Soc. Am. **109**, 2983–2987.

Jimenez, B. C. (**1987**). ''Acquisition of Spanish consonants in children aged 3–5,7,'' Lang. Speech Hear. Serv. Schools **18**, 357–363.

Kent, R. D., and Forner, L. L. (**1979**). ''Developmental study of vowel formant frequencies in an imitation task,'' J. Acoust. Soc. Am. **65**, 208–217.

Kent, R. D., and Vorperian, H. K. (**1995**). ''Development of the craniofacial–oral–laryngeal anatomy: A review,'' J. Medical Speech-Language Pathology **3**, 145–190.

Lehiste, I. (**1962**). *Acoustical Characteristics of Selected English Consonants* (Communication Sciences Laboratory, University of Michigan, Report No. 9, Ann Arbor, MI).

McGowan, R. S., and Nittrouer, S. (**1988**). ''Differences in fricative production between children and adults: Evidence from an acoustic analysis of /ʃ/ and /s/,'' J. Acoust. Soc. Am. **83**, 229–236.

Menyuk, P., and Anderson, S. (**1969**). ''Children's identification and reproductions of /w/, /r/, and /l/,'' J. Speech Hear. Res. **12**, 39–52.

Mochizuki, M. (**1981**). ''The identification of /r/ and /l/ in natural and synthesized speech,'' J. Phonetics **9**, 283–303.

Nittrouer, S. (**1995**). ''Children learn separate aspects of speech production at different rates: Evidence from spectral moments,'' J. Acoust. Soc. Am. **97**, 520–530.

Olive, J. P., Greenwood, A., and Coleman, J. (**1993**). *Acoustics of American English Speech* (Springer, New York).

Perkell, J. S., Boyce, S. E., and Stevens, K. N. (**1979**). ''Articulatory and acoustic correlates of the [s–ʃ] distinction,'' J. Acoust. Soc. Am. **65**, S24.

Sander, E. K. (**1972**). ''When are speech sounds learned?'' J. Speech Hear Disord. **37**, 55–63.

Shriberg, L., and Kent, R. (**1995**). *Clinical Phonetics* (Allyn and Bacon, Needham Heights, MA).

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., and Bird, A. (**1990**).

''The Iowa articulation norms and its Nebraska replication,'' J. Speech Hear Disord. **55**, 779–798.
Smit, A. B. (**1993**). ''Phonological error distributions in the Iowa–Nebraska articulation norms project: Consonant singletons,'' J. Speech Hear. Res. **36**, 533–547.
Stevens, K. N. (**1998**). *Acoustic Phonetics* (MIT Press, Cambridge, MA).
Stoel-Gammon, C. (**1985**). ''Phonetic inventories, 15–24 months: A longitudinal study,'' J. Speech Hear. Res. **28**, 505–512.
Strange, W., and Broen, P. A. (**1981**). ''The relationship between perception and production of /w/, /r/, and /l/ by three-year-old children,'' J. Exp. Child Psychol. **31**, 81–102.
Westbury, J. R. (**1994**). ''X-ray Microbeam Speech Production Database User's Handbook'' (University of Wisconsin, Madison, WI, unpublished).
Westbury, J. R., Hashi, M., and Lindstrom, M. J. (**1998**). ''Differences among speakers in lingual articulation for American English /ɹ/,'' Speech Commun. **26**, 203–226.

False beats in coupled piano string unisons

Brian Capleton
Royal National College, College Road, Hereford, HR1 1EB, United Kingdom

(Received 8 August 2003; revised 30 October 2003; accepted 4 November 2003)

The behavior of a unison pair of piano strings coupled by the soundboard bridge, when one string has localized anisotropy in the reactive part of the bridge admittance for a given partial frequency, can be investigated using a theoretical matrix description. The anisotropy can cause what in piano tuning terminology is referred to as ''false beating'' in a partial of the single string. A mathematical model can be used to illustrate how ''mistunings'' between the strings of the unison (measured when the strings are sounding in isolation from each other) may theoretically arise as a consequence of the normal practice in piano tuning, of eliminating or reducing audible beating in the unison when both strings are sounding. ''False beats'' in a single string partial can be ''inherited'' by a partial of the coupled unison's spectrum, and mistunings between the strings can eliminate or reduce the appearance of this inheritance. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1636461]

PACS numbers: 43.75.Mn, 43.40.Cw, 43.40.At [NHF] Pages: 885–892

I. INTRODUCTION

Modern pianos include ''bichord'' and ''trichord'' groups of two and three strings, respectively, each group serving for one note and all strings in the group being tuned to unison. The preference of auditors for unisons with ''mistunings'' between the strings of the group has been the subject of investigation by Kirk,[1] and acoustical reasons for mistunings being present as a consequence of expert tuning have been shown by Weinreich.[2] The term mistuning is defined as a frequency difference between partials with the same partial number, one for each string, when the strings are sounded in isolation from each other. Mistuning does not necessarily imply an ''out of tune'' condition of the unison with all strings sounding. Reasons for the presence of mistuning have been shown by Weinreich to include the fact that partial decay rate is functionally related to it. This provides the opportunity in tuning to compensate for the effects of ''irregularities'' in the piano hammer strike on the unison tone and decay.

It is proposed here, using a theoretical model, that a further reason for mistunings being present as a result of expert tuning may arise where the reactive part of transverse bridge boundary admittance is anisotropic, leading to ''false beats''[3] in audible partial decay. The term false beat[4] (''falseness'') is a standard part of piano tuning terminology, and refers to a beat pattern that is often present in the decay curve of one or more audible partial tones[5] of a single string, indicating that the audible partial is not radiated from a single normal mode, but from a mode *pair* close in frequency. Such a pair would occur within the mode frequency series that is otherwise approximately harmonic (except for some inharmonic dispersion due to string stiffness). It is proposed that the latter mistuning may occur where the artist tuner aims to eliminate beating as much as possible from the unison, which is standard tuning procedure.[6] This proposition is applicable even where the false beat rate in the isolated string may be so slow that the tuner may not necessarily have even regarded the string as ''false.''

Hundley, Benioff, and Martin reported that the vibration of a single string at one partial frequency ''does not typically exhibit a multiple decay rate,''[7] but Weinreich's recordings of a single string through a microphone placed near the piano[8] reveal a dual decay rate. Schuck and Young,[9] and Wolf and Sette[10] have also recorded beat patterns in the decays of various single-string partials. In general, single-string partials may decay with or without the presence of a regular beat pattern.

It is recognized that horizontal[11] and vertical string motions exhibit different decay rates, motion in the ''horizontal'' direction decaying more slowly. However, the hammer strike is ''vertical,''[12] and horizontal motion greater than that provided by the initial motion cannot establish itself unless normal modes exist whose mode coordinates are neither horizontal nor vertical. In either vertical or horizontal axes general motion following a hammer strike will then be a superposition of the normal modes. For slightly differing normal-mode frequencies, the superposition would be expected to be a decaying beat pattern. Very slow beats oscillating between horizontal and vertical directions have been observed in bass piano strings by Tanaka, Nagai, and Mizutani.[13]

II. "FALSE BEATS"

A. Practical effects of false beats

''False beats'' are a hazard in the attempt to eliminate beating in a unison or octave, or the attempt to set a specific beat rate in a tempered interval, because their beat rate is often inherited by the spectrum of the tuned interval, and often remains present no matter what tuning condition between the strings is selected. In tuner's terminology, a false beat cannot necessarily be ''tuned out,'' so it may frustrate any attempt to ''tune out'' beating altogether from the unison.

However, excellent artist tuners also know that often any such ''inherited'' beat in the unison pair (or indeed trichord) may be ''hidden'' by judicious fine-tuning adjustment, so that it is much less noticeable than the false beat in the single string. The technique of ''hiding'' falseness results in

changes to the audible decay patterns of false unison partials, and in particular the reduction of a false beat's audible amplitude. In some cases, after tuning a unison to ''hide'' slow falseness, the partial decay of the unison pair may even appear to be ''beatless'' when both strings are sounding. Despite this, the false string, if sounded on its own by muting the other strings with a wedge, still exhibits the false beat. This ability to hide or reduce the appearance of false beats is an essential part of the art of fine tuning.

B. The term "false" and incidence of falseness

The term ''false'' may have been ''traditionally'' used by many tuners to refer to a relatively fast-beating lower partial, and to imply a kind of ''fault.'' In fact, as long as the term false is used to refer to any discrete, audible partial that exhibits a beat or irregular pattern in its decay, then the term false can include all partials whose false ''beat rate'' may be very slow indeed—too slow for most tuners to call the string false, but still sufficient to alter the decay curve within the audible decay time. In these terms, false behavior may be thought of as relatively common, rather than occasional.

III. SINGLE STRING TRANSVERSE MOTION IN TWO PLANES

A. Normal modes

In a single plane, small-amplitude normal-mode motion in y of an ideal string, tensioned between rigid boundaries, with speaking length along the x axis, can be written in the form

$$y_n(x,t) = \mathrm{Re}\left[C_n \exp(in\omega_0 t) \sin\left(\frac{\pi n x}{l} \right) \right] \quad n = 1,2,...,\infty, \qquad (1)$$

where $\omega_0 = \pi c/l$ is the fundamental angular frequency, c is the phase velocity, and l is the speaking length. C_n can be regarded here as containing both amplitude and phase information. The series for n constitutes the mode numbers of modes with the harmonic series of frequencies $n\,\omega_0 = \omega_n$, determined by the phase velocity and the speaking length. In the case of a real piano string the mode frequencies would be subject to dispersion due to string stiffness (inharmonicity), but this will be of no consequence in the following discussion.

Replacing one of the rigid boundaries by a bridge boundary capable of small motion results in a change to the frequency ω_n, given by[14]

$$\delta\omega_n = iZ_n Y_n \omega_n / \pi, \qquad (2)$$

where ω_n is the frequency for a rigid boundary, Z_n is the wave impedance of the string, and Y_n is the complex bridge admittance. In the case where both the string and the movable bridge boundary are capable of transverse motion in two planes, the bridge boundary admittance can be specified as a complex matrix and the right-hand side of Eq. (2) can consequently also be evaluated as a complex matrix. Any elliptically polarized mode (when the real parts of the resulting eigenfrequencies are equal) will resolve into two modes linearly polarized with directions and phases determined by the bridge admittance matrix. However, the eigenfrequencies can also differ in their real parts, replacing one ω_n with two frequencies (in general, the transverse motion of the bridge will then describe a parametric locus). The former case corresponds to bridge admittance that in its reactive part is isotropic around x, and the latter case to anisotropic reactive bridge admittance.

In the discussion that follows, the *partial number* $j = 1,2,...,\infty$ corresponds to the series of n values, but labels motion that may contain two eigenfrequencies (in place of a single frequency ω_n) as a result of transverse anisotropy in the reactive part of the bridge admittance. These will be close compared to the separation of frequencies between two adjacent members of the harmonic series (in which the ω_n values fall), and can be responsible for the perceived audible false beat.

B. Parameter magnitudes

Using the grand piano arrangement in which the strike axis can be referred to as ''vertical,'' it is reasonable to assume that the decay rate in the horizontal axis will be about a quarter of that in the vertical axis. It is the ratio of these rates that affects the generic nature of the decay curve shape, and this ratio accords reasonably well with ratios in the ''vertical'' and ''horizontal'' directions that have been empirically recorded for single strings in grand pianos, by Hundley, Benioff, and Martin,[15] and Weinreich.[16] It might have been expected from the structural form of the bridge–soundboard system, that the reactive part of the bridge admittance would show a similar magnitude of variation between vertical and horizontal directions. However, Weinreich shows that the reactive part is approximately the same for both directions, and is only of comparable magnitude to the resistive part in the vertical direction.[17] Experimental results conclude that the angular variation of the reactive part of the bridge admittance is at least a factor of 10 smaller than the variation in the resistive part.[18] In the model proposed here, an almost isotropic reactive admittance would correspond to a situation in which the string would not exhibit a significant false beat at the frequency measured. Weinreich's investigations suggest a vertical bridge admittance of $Y \approx 10^{-3}$ s/Kg and a characteristic impedance of about 2 Kg/s for a middle octave string,[19] so from Eq. (2), the difference between eigenfrequencies when the reactive admittances are in, for example, the ratio 1:0.925, would for the fundamental of middle C at 261.65 Hz, produce less than 1 beat per 10 seconds. The estimated order of anisotropy in the reactive part of the bridge admittance, necessary to produce a beat frequency of at least 1 Hz, would be expected to be in the order of 6:1. For strings higher in the compass, the bridge admittance would be expected from design principles to be smaller relative to the string's characteristic impedance, but the fundamental frequencies are much greater. As a result, the anisotropy at the higher treble string boundaries necessary to produce a beat frequency of at least 1 Hz would be expected to be much smaller.

Although Weinreich's results showed a variation of this magnitude not to be present, Sec. IV below discusses how local physical boundary conditions might nevertheless cause large localized anisotropy in the reactive part of the bridge

admittance. Rather than being considered an anomaly, the local effects of the individual bridge (side draft) pin, for example, might actually explain how the reactive admittance could in other cases be almost isotropic, when this would not normally be expected from the overall structural design of the bridge–soundboard system.

IV. DISCUSSION OF PHYSICAL CAUSES

It is possible that factors other than anisotropic boundary conditions, such as nonuniformity in the string, or torsional stress, may cause false beats, but the presence of locally anisotropic boundary conditions may remain as an important cause in the case of piano strings *in situ*, where strings are of the highest quality and there are no stringing faults.

The soundboard bridge (''long bridge'') end of the speaking length of the string, terminates in an arrangement where the string physically continues across the bridge surface (about 2 cm) but is drafted sideways a few degrees past the first side draft pin (''bridge pin''). The pin is angled into the bridge surface to resist the side-draft force of the string, and to ensure a good contact between the string and the flat bridge surface. The pin itself may not be acoustically rigid. The small dimensions of the pin, coupled with the high tension in the string it drafts sideways, leads to a significant pressure in the wood around the pin hole, which on some older instruments develops very small splits. Where there is an actual fault such as loss of downbearing,[20] the string may not fully contact the bridge surface at the pin, and may even pass the pin fractionally above the surface of the bridge. The pin then effectively acts as a cantilevered support of the string, its own elastic properties being significant as a possible cause of anisotropic boundary conditions. For an arbitrary speaking length of 1 m, and a string diameter of 1 mm, at middle C (261 Hz) a speaking length variation of 0.1 mm between two normal planes will only account for a frequency difference between the two planes of around 1/40th Hz in the fundamental. However, for the top C88 with a speaking length say, 0.08 that of middle C, the same variation in speaking length can account for a frequency difference of around 5 Hz in the fundamental. The magnitude of variation in *localized* conditions (confined to the influence of the bridge pin) may be similar for all strings, resulting in an expected greater frequency difference in the treble and higher treble strings. In the present model, variation in speaking length l measured on different planes can also be included as part of the anisotropy in the boundary admittance for a constant speaking length, given that the physical effect of reactive boundary admittance can be conventionally considered as a lengthening or shortening of the wavelength on the string, due to a shift in the position of the boundary node.

V. COUPLED STRINGS: HIDING OR DISGUISING FALSENESS

A. Mathematical model

Weinreich has shown that in the fine-tuning region, the bridge coupling between the strings of a unison pair leads to behavior significantly different from that which would be expected for the uncoupled system.[2] It is possible to propose a further model representing a situation where motion in two planes is accounted for, and one string has anisotropic localized boundary conditions at the bridge, leading to falseness. In this case the uncoupled system has potentially four oscillators corresponding to the two pairs of individual string normal modes (it cannot be assumed, however, that in the general case the normal modes of the coupled system have their mode axes aligned with those of the uncoupled system). Figure 1 represents the model adopted for the arrangement of oscillators in the uncoupled system, for strings 1 and 2.

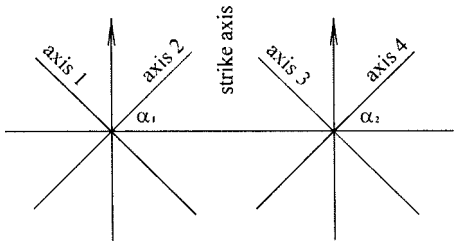


FIG. 1. Arrangement of uncoupled mode axes at the bridge boundary in the Eq. (6) model.

The bridge admittance in this model, in the absence of localized variations, is based on the parameters indicated by Weinreich's results and as discussed in Sec. III B above, i.e., its reactive part is isotropic around the string's speaking length axis, while its resistive part is the same magnitude as the reactive in the vertical axis, but a quarter of this in the horizontal.

For one partial j the uncoupled system thus comprises four oscillators, each oscillator k having displacement q in the form

$$q_k = \mathrm{Re}(\psi_k). \tag{3}$$

The general equation of motion for the coupled system derived from the perturbation theory[21] is

$$\frac{d\Psi}{dt} = i\mathbf{\Omega}\Psi, \tag{4}$$

where Ψ is a vector of all the ψ_k's, and $\mathbf{\Omega}$ is a matrix whose elements kk' are the frequency perturbation on oscillator k by the coupling mechanism and oscillator k'. For a perturbation 2ε due to ''mistuning'' considered applied to one string only, the matrix extended from that used by Weinreich, is in the form

$$\mathbf{\Omega} = \begin{pmatrix} \gamma_{11}+L_1+2\varepsilon & 0 & \gamma_{13} & \gamma_{14} \\ 0 & \gamma_{22}+2\varepsilon & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33}+L_3 & 0 \\ \gamma_{41} & \gamma_{42} & 0 & \gamma_{44} \end{pmatrix}, \tag{5}$$

where complex γ_{mn} is the perturbation on oscillator m by the bridge and oscillator n, and L_n is the localized perturbation on axis n.

When $\alpha_1 = \alpha_2 = 0$ or 2π, and $L_1 = L_3 = 0$ (corresponding to isotropic reactive bridge admittance and only two normal-mode motions, aligned with the strike axis), the eigenfrequencies of $\mathbf{\Omega}$ are the same as those for the 2×2 matrix in the single plane model. In the latter, illustrated by Weinreich, the coupled system exhibits two modes, one ''symmetric,''

and one ''antisymmetric,'' aligned with the conventionally chosen vertical direction of the single-string oscillators. Effectively, in the symmetric mode, the two oscillators are ''in phase,'' and in the antisymmetric mode they are ''out of phase.'' The present model for two planes also yields modes that are *symmetric* and *antisymmetric*, but in this case now they are in orthogonal pairs. There is a large number of possible configurations of the parameters, and the proposed principle of hiding false beats, whose effectiveness does depend on the parameter configuration, can be illustrated through a simplified matrix putting $\alpha_1=\alpha_2=\pi/4$, so that

$$\mathbf{\Omega}=\begin{pmatrix} \gamma+L_1+2\varepsilon & 0 & \gamma & 0 \\ 0 & \gamma+2\varepsilon & 0 & \gamma \\ \gamma & 0 & \gamma+L_3 & 0 \\ 0 & \gamma & 0 & \gamma \end{pmatrix}. \tag{6}$$

Setting $\alpha_1=\alpha_2$ for the case where one string only is false means that the isolated motion of the other (nonfalse) string is still considered as two oscillators in the same two planes as the isolated false string's mode axes (even though it could be resolved to one plane). Using $\alpha=\pi/4$, the initial displacements of the component oscillators are also equal.

B. Hypothetical resistive bridge

In the hypothetical ''test'' case of a purely resistive bridge, both the real and imaginary parts of the eigenfrequencies of $\mathbf{\Omega}$ in Eq. (6), appear in two corresponding pairs. When there are no localized perturbations, the two pairs of curves for the real and imaginary parts of the eigenfrequencies—corresponding, respectively, to the coupled system's normal-mode frequencies and their decay rates—each coincide. The graphs then appear visually the same as those for the one plane model, as illustrated by Weinreich. When a localized perturbation is introduced, this has the effect of shifting one pair of eigenfrequencies (both the real and the imaginary) relative to both axes. Figures 2(A) and (B) show the effect of introducing a localized perturbation on axis 3, by putting $L_3=3$. The real frequencies of each mode pair coalesce across a ''central'' range of mistunings as they would in the one-plane model, but the coalescing range for one of the pairs is no longer centered on zero mistuning, as it is in the one-plane model. In the latter, mistunings in the range 0 to ± 1 (the single-string decay rate being equal to unity) produce no beat (there is only one real frequency for any mistuning value in this range) but there is a dual decay rate. In the two-plane model the number of differing real frequencies present is never less than two, which occurs when the mistuning is in the range 0.5–1.0. The resulting beat rate (as the difference in real frequencies) is constant throughout this range. In the middle of this range at a mistuning of 0.75, each real frequency has associated with it the same two different decay rates. Mistunings close to this point, rather than at zero mistuning, might be expected to yield a decay curve that could perhaps be accepted by the piano tuner as a substitute for a beatless decay, where the ideal (but in this case unattainable) aim is to eliminate beats or fluctuations altogether.

C. "Realistic" bridge

In the case of ''realistic'' bridge parameters with the reactive component present, the eigenfrequency pairs are again offset by the presence of a single localized perturbation, Figs. 2C and (D). If $L_1=L_3\neq 0$ (two equal perturbations) the decay rates (imaginary eigenfrequencies) then coincide, but the real eigenfrequency pairs are still offset in frequency, Fig. 2(E).

In the hypothetical example of a purely resistive bridge, at zero mistuning the antisymmetric modes have zero decay rates (if an antisymmetric mode is the only mode present, the bridge will not move), and the symmetric modes have twice the single-string decay rate in that axis direction. However, as in Weinreich's model, an antisymmetric mode will not be present at zero mistuning unless there is a difference in the initial conditions of the two component oscillators. Weinreich suggests how ''imperfections'' in the hammer could lead to the presence of the antisymmetric mode, from the initiation of the system. The present model takes into account mistuning as the only factor that introduces the antisymmetric modes. In the physical case imperfection in the hammer strike, including strike misalignment, especially when there are grooves worn in the hammer face, could also contribute to antisymmetric mode excitation.

For finding generic decay curve shape characteristics, the eigenfrequencies of $\mathbf{\Omega}(\varepsilon)$ are applied as perturbations to an arbitrary set of mode frequencies that are equal in their real parts.

Each mode p_m of the coupled system (where $m=1$ to 4) is written as

$$p_m(t,\varepsilon)=\mathrm{Re}[\exp\{i(\omega_0+v_{\mathbf{R}m}(\varepsilon)+v_{\mathbf{I}m}(\varepsilon))t\}], \tag{7}$$

where $v_{\mathbf{R}m}(\varepsilon)$ is the real part of eigenfrequency m and $v_{\mathbf{I}m}(\varepsilon)$ is its imaginary part. The unperturbed frequency ω_0 here can in principle be any value (since we are interested only in functional dependence on time and ε). However, in the present model, the resultant decay curves were plotted from the envelope obtained by sampling the actual waveform produced in the strike axis direction. Thus, ω_0 was included at a suitably high value (arbitrarily set at 440 Hz) compared to the sampling frequency.

Two orthogonal general motions, each superposed from one symmetric and one antisymmetric mode, are possible, and these are written in the model as

$$G_1(t,\varepsilon)=p_1v_{\mathbf{I}1}+p_2v_{\mathbf{I}2}, \quad G_2(t,\varepsilon)=p_3v_{\mathbf{I}3}+p_4v_{\mathbf{I}4}. \tag{8}$$

Here, each waveform p is multiplied by its decay rate. This is a device for generating a curve ''emulating'' the radiated sound intensity, which is assumed to correlate positively with the decay rate. Thus, at zero mistuning when the antisymmetric decay rate is zero, its initial amplitude is zero (corresponding to the hammer strike being ''perfectly'' equal on both strings, with no antisymmetric component present). Similarly, the symmetric decay rate will at zero mistuning approach double its value at larger mistunings, providing the louder, but faster decaying ''prompt sound.''

Setting the α's with equal values of 45 degrees produces a resultant waveform $C(t)$ in the ''strike'' axis direction, for any fixed mistuning value, in the form

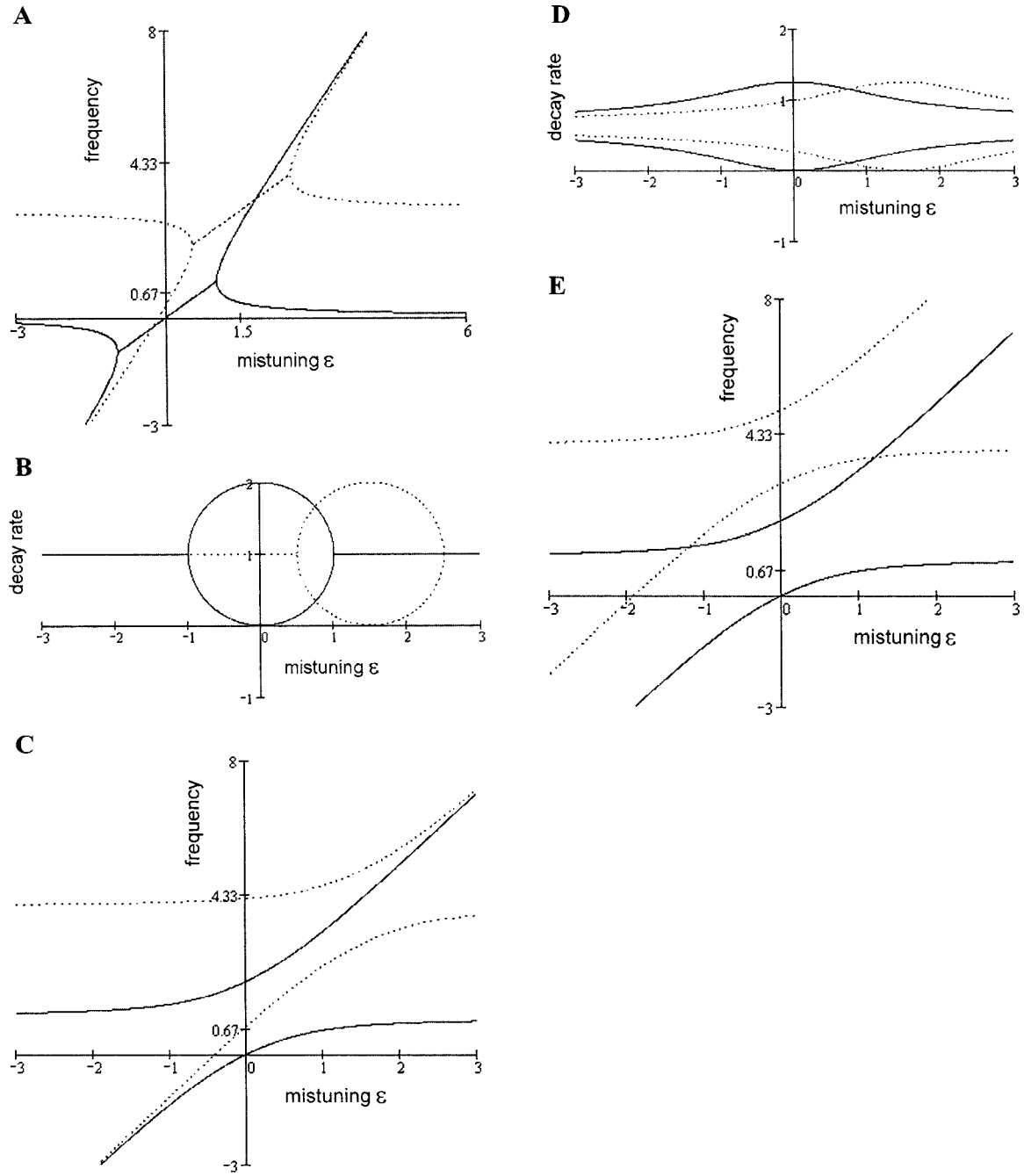


FIG. 2. Eigenfrequencies of $\mathbf{\Omega}$. (A) Real parts for a hypothetical purely resistive bridge when $L_3=3$ and $L_1=0$ corresponding to one false string with a false beat frequency of about 0.7 Hz. (B) Decay rates (imaginary part) of the same eigenfrequencies for the purely resistive bridge. (C) Real parts for a "realistic" reactive and resistive bridge when $L_3=3$ and $L_1=0$. (D) Decay rates for the reactive and resistive bridge of when $L_3=3$ and $L_1=0$. (E) Real parts when $L_3=L_1=3$, corresponding to two false strings.

$$C(t)=A_1G_1(t)\sin 45°+A_2G_2(t)\cos 45°. \tag{9}$$

For equal initial string displacements, the generic decay curve shape characteristics can be plotted using simply

$$C(t)=G_1(t)+G_2(t). \tag{10}$$

To generate a curve representing the overall generic decay pattern in the partial, $|C(t)|$ was sampled by integrating forward over 1/100th of a second (roughly 4 cycles) at each sample point, at 20 samples per second, for 5 s. This produces a set of samples E_n for $n=1,2,...,100$ for the decay curve. Thus, for unit unperturbed decay rate the graphs of y_n (in dB) are for

$$y_n=10\left[\log\left(E_n\exp-\frac{n}{20}\right)-\log\left(E_1\exp-\frac{1}{20}\right)\right], \tag{11}$$

where the horizontal axis n represents units of 1/20th s. The value of ε in $G_1(t,\varepsilon)$ and $G_2(t,\varepsilon)$ in the generation of all the graphs is given using $\varepsilon=\pi M$, where M is the mistuning in

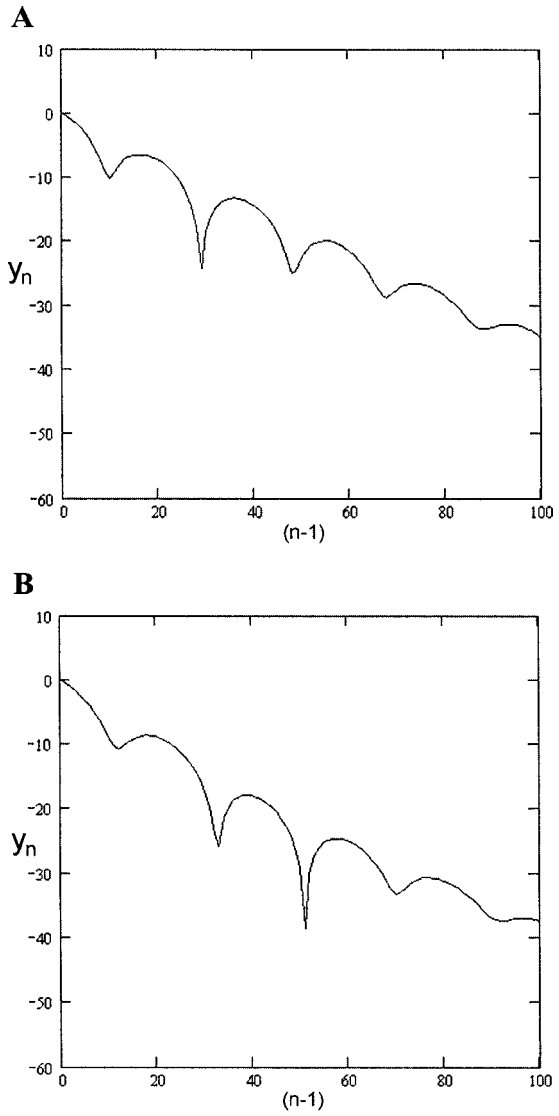


FIG. 3. Generic decay curves—beats due to mistuning and false beat inheritance. (A) Theoretical generic decay pattern of one partial *j* for the unison when neither string is false, and there is a mistuning of 1 Hz at partial *j*. *n* corresponds to 1/20th second. (B) Theoretical generic decay pattern of one partial *j* for the unison when one string has a false beat rate of 1 Hz at partial *j*, and there is no mistuning between the strings at partial *j*.

Hz,[22] for the model's theoretical single string partial *j* with unit decay rate.

When $L_1=L_3=0$ corresponding to a unison in which neither string is false (for this partial number), mistuning between the strings produces the expected decaying ''heterodyne beat-like'' pattern. Figure 3(A) shows the pattern for a mistuning of 1 Hz.

It was shown from Eq. (2), in Sec. III B above, that the local perturbation variation necessary to produce a false beat rate of 1 Hz would need to be in the order of 6:1. Putting L_1 or L_3 in Eq. (6) equal to 6, with zero mistuning, produces the curve Fig. 3(B), in which the beat from the false string is inherited by the decay of the coupled unison.

For a sufficiently small false beat rate, there is an ''optimum range'' of mistuning for which the false beat pattern is not ''obviously'' inherited by the unison decay. In this particular model the largest false beat rate for which such an optimum range can be found is approximately 0.7 Hz, for which L_1 or L_3 will equal 3 [Fig. 4(A)]. When this false beat rate is present, with increasing mistuning, the decay pattern temporarily becomes *less* ''turbulent'' [Fig. 4(B)], and more ''quasiexponential,'' up to mistuning values of around 0.3–0.4 Hz (which would produce a beat rate of around 1 beat every 2 or 3 s, in the absence of any falseness or bridge coupling). Figure 4(C) shows a decay in which the original false beat rate is present but only just discernible, having a greatly reduced beat amplitude. For mistunings greater than this, ''turbulence'' in the decay begins to increase [Fig. 4(D)] until at around 2-Hz mistuning, a beat rate equivalent to the mistuning begins to assert itself in addition to the false beat [Fig. 4(E)]. For the faster false beat rate ($L_3=6$) of 1 Hz no such position giving a quasiexponential or ''beatless'' decay or dual decay, can be reached.

VI. CONCLUDING REMARKS

The general principle illustrated is that false beats in the decay curve of a partial of a single piano string, caused by anisotropy in the bridge admittance, can be inherited in the decay curve of the unison string pair (and by implication, the trichord group). However, the beat in the unison group partial may be hidden at certain values of mistuning between the strings, if the single-string false beat rate is not too fast compared to the decay rate. No general algebraic treatment of the functional relationship between parameters has been developed here, but such an understanding is called for. The model presented here is limited; its purpose is to engender further interest in false beat phenomena and its effect on unison tuning. In particular, the concept of hiding false beats through mistuning deserves further investigation.

Because expert piano tuners adjusting the tone of the unison cannot do so by listening merely to the fundamental (since the tone is not only dependent on the fundamental), it is proposed here that reference should be made to a subset of the audible partials in the unison spectrum that constitutes a set of ''significant partials'' for tuning purposes. It can be speculated that what constitutes significant may not be entirely objective, and that hearing psychology may play a part. Even though a beat inherited at zero mistuning by the unison spectrum may subsequently be hidden by introducing a certain small degree of mistuning between the strings, the required mistuning at a given partial number may or may not be suitable (depending on the many parameters) to simultaneously hide falseness at *another* partial number *j*. Thus, the setting of mistuning between the strings where falseness is present may sometimes be to achieve, at best, a minimum of beating in one or more significant partials, where false beat behavior prevents beat elimination. For sufficiently slow false beat rates, it can be proposed that eliminating beating altogether from the significant partials may still be possible.

From a generalized pure mathematical description for the formation of close-frequency mode pairs in piano strings, it may be possible to derive falseness due to boundary admittance anisotropy, as a specific case. Experienced piano tuners are familiar with a very large range of varying tuning characteristics in pianos, indicating that a general empirical understanding demands investigation of a large number of

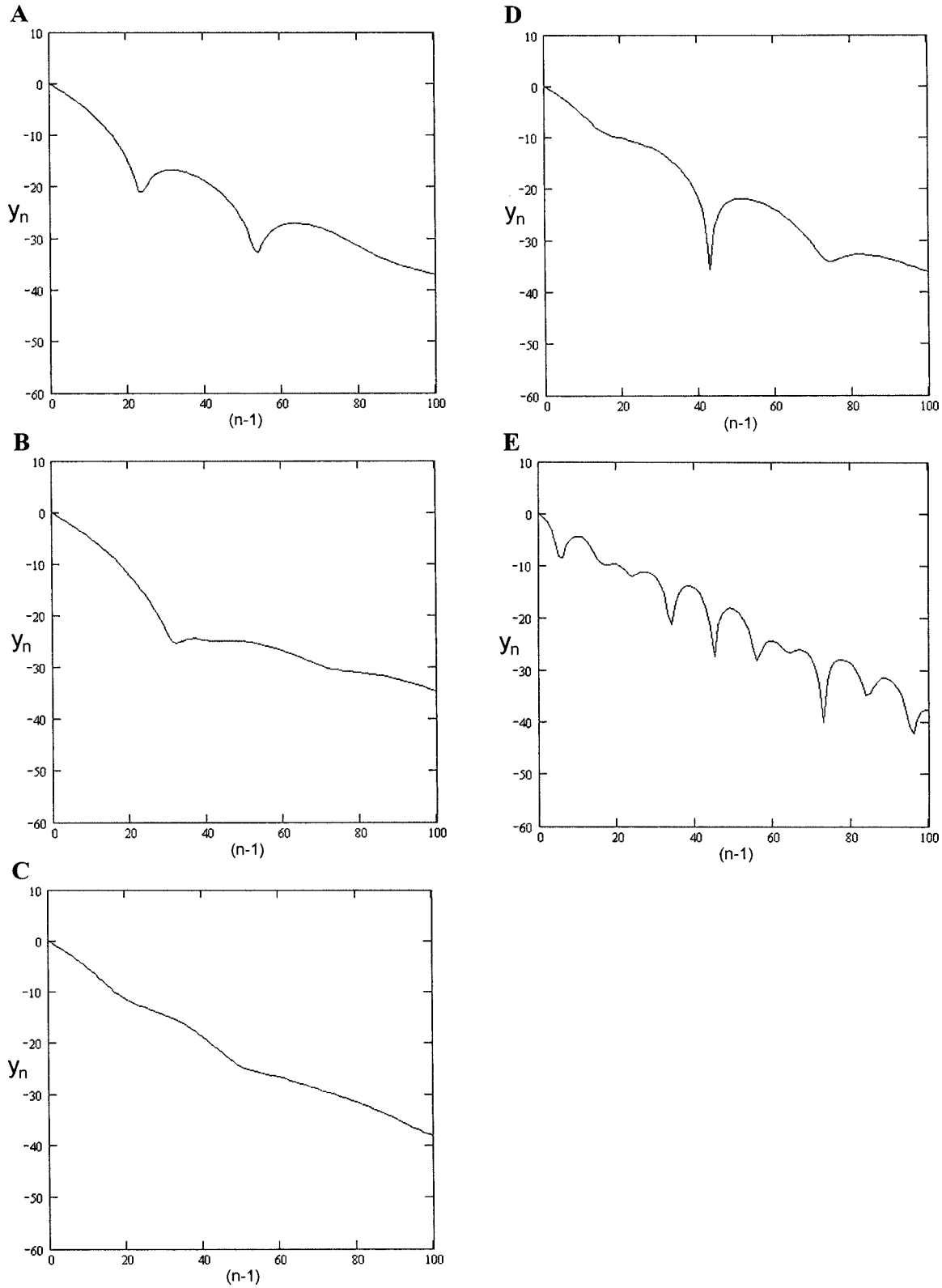


FIG. 4. Generic decay curve patterns of one partial *j* for a unison with one false string, with increasing mistuning. (A) One string has a false beat rate of 0.7 Hz at partial *j*, and there is no mistuning between the strings at partial *j*. *n* corresponds to 1/20th second. (B) Mistuning between the strings at partial *j* increased to 0.2 Hz. (C) Mistuning increased to 0.4 Hz. This is probably the ''optimum'' mistuning for achieving a ''quasiexponential'' or ''beatless'' decay. (D) Mistuning increased to 0.6 Hz. (E) Mistuning increased to 2 Hz. The beat pattern is now a mixture of both the false beat and a beat due to the mistuning.

strings over a substantial cross section of instruments by different manufacturers. In relation to the present theoretical model, such investigations would require the search for transverse normal-mode axes, and also possibly measurements of audible partial frequency variation as a function of time. Empirical disproving of any part of the present propositions may also suggest or reveal other mechanisms currently unrecognized.

ACKNOWLEDGMENT

Gabriel Weinreich is gratefully acknowledged for his valuable comments on the first version of this paper.

[1] R. E. Kirk, ''Tuning preferences for piano unison groups,'' J. Acoust. Soc. Am. **31**, 1644–1648 (1959).
[2] G. Weinreich, ''Coupled piano strings,'' J. Acoust. Soc. Am. **62**, 1474–1484 (1977).
[3] See, for example, W. Braid White, *Piano Tuning and Allied Arts* (1917, 14th reprint 1972), p. 106. The description includes some unsupported assumptions about the physics of falseness.
[4] In older texts this is sometimes called a ''false wave.'' See, for example, J. Cree Fischer, *Piano Tuning/Regulating and Repairing* (1907), reprinted as *Piano Tuning. A Simple and Accurate Method for Amateurs* (Dover, New York, 1975), pp. 160–161.
[5] The component partials of the decay tone (in an approximately harmonic frequency series) can with training be heard in practice, as discrete audible tones.
[6] The idea that false beats can be ''neutralized'' by tuning the false string ''slightly off'' from the other strings in the group was suggested by W. Braid White, *Piano Tuning and Allied Arts* (1917, 14th reprint 1972), p. 106.
[7] T. C. Hundley, H. Benioff, and D. W. Martin, ''Factors contributing to the multiple rate of piano tone decay,'' J. Acoust. Soc. Am. **64**(5), 1303–1309 (1978). See, in particular, p. 1306.
[8] G. Weinreich, in Ref. 2, p. 1479.
[9] O. H. Schuck and R. W. Young, J. Acoust. Soc. Am. **15**(1), 1–11 (1943).
[10] S. K. Wolf and W. Sette, J. Acoust. Soc. Am. **6**, 160–168 (1935).
[11] ''Horizontal'' and ''vertical'' refer, respectively, to the direction parallel and perpendicular to the surface of a grand piano soundboard.
[12] The grand piano arrangement is adopted as a standard reference.
[13] H. Tanaka, K. Mizutani, and K. Nagai, ''Experimental analysis of two-dimensional vibration of a piano string measured with an optical device,'' J. Acoust. Soc. Am. **105**(2), 1181 (1999), and ''Two-dimensional motion of a single piano string,'' *ibid.*, **100**(4), 2843 (1996).
[14] See P. M. Morse, *Vibration and Sound* (New York, 1948), Sec. III, p. 13.
[15] T. C. Hundley, H. Benioff, and D. W. Martin, ''Factors contributing to the multiple rate of piano tone decay,'' J. Acoust. Soc. Am. **64**(5), 1303–1309 (1978).
[16] G. Weinreich, in Ref. 2, p. 1480.
[17] G. Weinreich, in Ref. 2, p. 1480.
[18] G. Weinreich, in Ref. 2, p. 1481.
[19] G. Weinreich, in Ref. 2, p. 1482.
[20] From the soundboard bridge, the string angles downward towards the hitch pin on the iron frame. This angle is the ''downbearing angle.'' The downbearing angle keeps the string firmly on the bridge surface. The bridge is ''crowned'' with a two-dimensional arch supporting the total downbearing force of the strings against it. Thus, the entire soundboard bridge and strings system acts as an elastic, acoustical structure.
[21] G. Weinreich, in Ref. 2, pp. 1474–1475.
[22] G. Weinreich, in Ref. 2, p. 1475.

Wave equations, dispersion relations, and van Hove singularities for applications of doublet mechanics to ultrasound propagation in bio- and nanomaterials

Junru Wu[a] and Christopher Layman
Department of Physics, University of Vermont, Burlington, Vermont 05405

Jun Liu
Biomedical Engineering Center, The Ohio State University, Columbus, Ohio 43210

(Received 25 April 2003; revised 10 October 2003; accepted 24 November 2003)

A fundamental mathematical framework for applications of Doublet Mechanics to ultrasound propagation in a discrete material is introduced. A multiscale wave equation, dispersion relation for longitudinal waves, and shear waves are derived. The van Hove singularities and corresponding highest frequency limits for the Mth-order wave equations of longitudinal and shear waves are determined for a widely used microbundle structure. Doublet Mechanics is applied to soft tissue and low-density polyethylene. The experimental dispersion data for soft tissue and low-density polyethylene are compared with results predicted by Doublet Mechanics and an attenuation model based on a Kramers–Kronig relation in classical continuum mechanics. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1642620]

PACS numbers: 43.80.Cs, 43.35.Cg [FD] Pages: 893–900

LIST OF SYMBOLS

λ	Lame's constant
$\mathbf{A}_{\alpha\beta}$	microlevel elastic moduli matrix
$\mathbf{A}_{12}$	$= \lambda + \mu$
$\mathbf{A}_{44}$	$= 4\mu$
$\mathbf{p}_\alpha$	microstress
α,β	indices for nodes
τ_α	directional cosine unit vector
ζ_α	inter-nodal distance after elongation
$\mathbf{C_{ijkl}}$	$= \Sigma_{\alpha=1}^{n}\Sigma_{\beta=1}^{n}\mathbf{A}_{\alpha\beta}\tau_{\alpha i}^{0}\tau_{\alpha j}^{0}\tau_{\beta k}^{0}\tau_{\beta l}^{0}$
$\mathbf{C_{S,M}}$	Mth-order shear phase velocity
ω	angular frequency
$\mathbf{f}$	frequency
μ	Lame's constant
$\mathbf{A}_{11}$	$= \lambda + 4\mu$
$\mathbf{A}_{15}$	$= -2\mu$
σ_{ij}	macrostress
ρ	mass density
M	order of Taylor's expansion
η_α	internodal distance
ε_α	elongation (fractional change of internodal distance)
$\mathbf{C}_{L,M}$	Mth-order longitudinal phase velocity
$\mathbf{k}$	wave number
λ_0	wavelength
$\mathbf{u}$	displacement of a node

I. INTRODUCTION

Doublet Mechanics (also called Nanomechanics) is a new elastic theory dealing with materials that possess discrete microstructures (Granik, 1997) such as some bio- and nanomaterials. Classical continuum mechanics (CCM) assumes that materials are continuous; there is no scaling factor that can be used to relate to the discrete property of those materials.

In Doublet Mechanics (DM), the fundamental microscopic unit of a material can be represented by a bundle of spheres or nodes separated by a finite distance (called internodal distance) (Ferrari *et al.*, 1997). Nodes in DM could be cells in biological cases or molecules of a solid. Nodes are either in contact with each other or separated by adhesive layers. In DM, the mechanical strain developed in a bundle due to its distortion is a function of the internodal distance. Depending on the order and the discreteness of the system, it can be expressed as a multiscale Mth-order Taylor expansion of the internodal distances at its equilibrium configuration. Readers should not relate the order (M) of the Taylor expansion to the degree of accuracy of this presentation in a common mathematical sense. In fact, the order of M represents the degree of the discreteness of a system; the higher the M order a Taylor expansion includes, the more discrete nature of the system DM represents (Ferrari *et al.*, 1997). When $M=1$, i.e., only the first order of the Taylor expansion is used; it reduces to the classical continuum case, i.e., the system described by DM becomes a system that does not have a discrete nature at all. The results obtained from DM completely agree with those predicted by CCM. When $M=\infty$, on the other hand, it describes a single crystal that has a regular and periodic lattice structure. When $\infty > M > 1$, the discreteness of systems represented by DM are between continuous (nongranular) and periodically discrete. A recent success of DM in an ultrasonic application to the characterization of elastic properties of tissue at relatively low frequency (<10 MHz) shows that DM is quite versatile and powerful in the description of wave propagation in partially discrete and granular materials such as biological soft tissue (Liu and Fer-

[a]Author to whom correspondence should be sent. Electronic mail: Jun-Ru.Wu@uvm.edu

rari, 2002). In this application, the Taylor expansion of $M=2$ was applied and cells were treated as nodes. The fact that Liu and Ferrari (2002) were able to use $M=2$ to describe soft tissue successfully indicates that the discreteness of soft tissue is much less than a crystal and more close to a nondiscrete continuum medium. The potential of DM in applications such as ultrasound imaging and nondestructive evaluation of granular materials is enormous, and research in those fields is just in its infant age.

Soft tissue (Bamber, 1986; Kremkau *et al.*, 1981) and several polymer materials (Wu, 1996; He, 2000; Szabo and Wu, 2000) were found to be dispersive in the medical ultrasound frequency range (1–10 MHz); their sound speeds were an increasing function of frequency. Efforts were made to explain the dispersion in terms of attenuation (O'Donnel *et al.*, 1981; Szabo, 1994; Szabo and Wu, 2000). A general procedure suggested by Szabo and Wu (2000) may be given as follows. First, the nondissipative Hooke's law that relates stress and strain was modified to include dissipation. Consequently, a nonloss wave equation was also modified to include dissipation. Thus, the wave number of a general plane-wave solution of the modified wave equation became a complex number; its imaginary part corresponded to attenuation. Based on a Kramers–Kronig relation, an equation that relates dispersion (the excess wave velocity) and attenuation was derived (Szabo and Wu, 2000). Therefore, if attenuation as a function of frequency is measured, the corresponding dispersion could be calculated, or *vice versa*.

It was also believed that the frequency-dependent behavior of ultrasound velocity might be due to a higher-order structure of the biological soft tissue and polymers (Bamber, 1986). In other words, the dispersion could arise from the medium has internal structure and degrees of freedom (Auld, 1990) other than attenuation.

In this paper, an Mth-order lossless wave equation is derived, and the dispersion relation for various orders of M are calculated for a commonly used microconfiguration that will be described below. Our numerical results show that the dispersion behaves quite differently for different orders of M. It reflects the richness and flexibility of DM, i.e., by choosing the appropriate order of M, DM may accommodate very different discrete physical systems. Particularly, for $M=2$, we were able to show an upswing dispersion, i.e., the phase velocity is an increasing function of frequency without considering attenuation. Experimental data of attenuation for a soft tissue and polymers were fitted to a power law of frequency, dispersion due to attenuation was derived by applying an equation based on a Kramers–Kronig relation for the soft tissue and polymers. It was found that at low frequency, dispersion due to attenuation is much greater than that due to its discreteness predicted by DM. As frequency increases (the wavelength decreases), especially, when the wavelength is comparable to the internodal distance, dispersion generated by the discrete nature of the sample predicted by DM becomes a dominant factor.

A. The framework of DM

The fundamental microstructure in DM is called a doublet bundle that contains a node and its α neighboring nodes separated by their internode distance of η_α. Figure 1 shows a fundamental microstructure of a doublet bundle that is commonly used in DM (Ferrari, 2000). A node of **a** (Fig. 1) has $\alpha=\mathbf{6}$ neighboring nodes. Their orientation can be expressed by six directional cosine unit vectors $\tau_\alpha(\alpha=1,2,...,6)$: $\tau_1=(1,0,0)$, $\tau_2=(0,1,0)$, $\tau_3=(0,0,1)$, $\tau_4=(0,1/\sqrt{2},1/\sqrt{2})$, $\tau_5=(1/\sqrt{2},0,1/\sqrt{2})$, and $\tau_6=(1/\sqrt{2},1/\sqrt{2},0)$ in Cartesian coordinates. A tensor consists of these six unit vectors may be expressed by a 6×3 matrix τ as below:

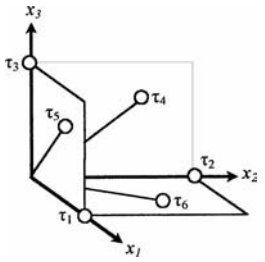


FIG. 1. An illustration of Cartesian coordinates used in the description of six nodes of a doublet bundle: $\tau_1=(1,0,0)$, $\tau_2=(0,1,0)$, $\tau_3=(0,0,1)$, $\tau_4=(0,1/\sqrt{2},1/\sqrt{2})$, $\tau_5=(1/\sqrt{2},0,1/\sqrt{2})$, and $\tau_6=(1/\sqrt{2},1/\sqrt{2},0)$.

$$\tau=\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \mathbf{0} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \mathbf{0} \end{bmatrix}. \tag{1}$$

The reason that this particular doublet bundle was chosen is for mathematical convenience (Ferrari, 2003).

Figure 2 illustrates a doublet of nodes **a** and **b** separated by η_α. After a translational distortion, they move to $\mathbf{a}'$ and $\mathbf{b}'$, respectively. The elongation (fractional change of internodal distance) of this arbitrarily chosen doublet (somewhat equivalent to strain in CCM) may be written as

$$\varepsilon_\alpha=\frac{\zeta_\alpha-\eta_\alpha}{\eta_\alpha}, \tag{2}$$

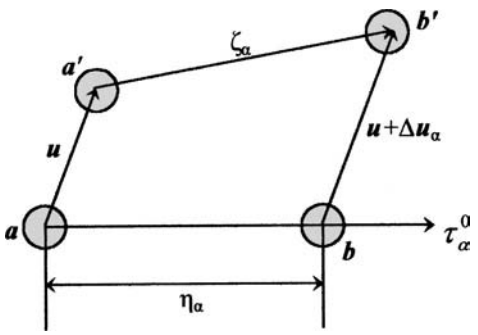


FIG. 2. An illustration of distortion between a doublet.

where ζ_α shown in Fig. 2 can be expressed by a Taylor expansion in Cartesian coordinates (Ferrari *et al.*, 1997) as

$$\zeta_\alpha \approx \eta_\alpha + \tau^0_\alpha \cdot \Delta \mathbf{u}_\alpha, \tag{3}$$

where

$$\Delta \mathbf{u}_\alpha = \sum_{\chi=1}^{\mathbf{M}} \frac{(\boldsymbol{\eta}_\alpha)^\chi}{\boldsymbol{\chi}!} \boldsymbol{\tau}^0_{\boldsymbol{\alpha}\mathbf{k}_1} \cdots \boldsymbol{\tau}^0_{\boldsymbol{\alpha}\mathbf{k}\chi} \frac{\partial^\chi \mathbf{u}_i}{\partial \mathbf{x}_{\mathbf{k}_1} \cdots \partial \mathbf{x}_{\mathbf{k}\chi}}\bigg|_{\mathbf{x}=\mathbf{x}^0}. \tag{4}$$

Therefore the elongation ε_α is given by

$$\varepsilon_\alpha = \tau^0_{\alpha i} \sum_{\chi=1}^{\mathbf{M}} \frac{(\eta_\alpha)^{\chi-1}}{\chi!} \tau^0_{\alpha k_1} \cdots \tau^0_{\alpha k\chi} \frac{\partial^\chi \mathbf{u}_i}{\partial \mathbf{x}_{\mathbf{k}_1} \cdots \partial \mathbf{x}_{\mathbf{k}\chi}}\bigg|_{\mathbf{x}=\mathbf{x}^0}. \tag{5}$$

The equality $\mathbf{x}=\mathbf{x}^0$ in Eq. (5) indicates that the differentiation in that equation is evaluated at the coordinates of the doublet nodes at equilibrium, i.e., $\mathbf{x}^0_{\mathbf{k}1},\ldots,\mathbf{x}^0_{\mathbf{k}\boldsymbol{\chi}}$. Each subscript of the set $\mathbf{k}_1,\ldots,\mathbf{k}_{\boldsymbol{\chi}}$ runs through the integers 1, 2, 3 corresponding to three-dimensional (3-D) Cartesian coordinates. The Mth-order of Eq. (5) may be calculated by keeping M terms, for example, for $M=1$,

$$\varepsilon_\alpha = \tau^0_{\alpha\mathbf{i}} \tau^0_{\alpha\mathbf{j}} \frac{\partial \mathbf{u}_\mathbf{i}}{\partial \mathbf{x}_\mathbf{j}}\bigg|_{\mathbf{x}=\mathbf{x}^0} = \tau^0_{\alpha\mathbf{i}} \tau^0_{\alpha\mathbf{j}} \varepsilon_{\mathbf{ij}}\big|_{\mathbf{x}=\mathbf{x}^0}, \tag{6}$$

where

$$\varepsilon_{\mathbf{ij}} = \frac{1}{2}\left(\frac{\partial \mathbf{u}_\mathbf{i}}{\partial \mathbf{x}_\mathbf{j}} + \frac{\partial \mathbf{u}_\mathbf{j}}{\partial \mathbf{x}_\mathbf{i}}\right) \tag{7}$$

Here $\varepsilon_{\mathbf{ij}}$ corresponds to strain in the CCM case (Auld, 1990).

For $M=2$,

$$\varepsilon_\alpha = \left(\tau^0_{\alpha\mathbf{i}} \tau^0_{\alpha\mathbf{j}} \frac{\partial \mathbf{u}_\mathbf{i}}{\partial \mathbf{x}_\mathbf{j}} + \tau^0_{\alpha\mathbf{i}} \frac{\eta_\alpha}{2} \tau^0_{\alpha\mathbf{j}} \tau^0_{\alpha\mathbf{k}} \frac{\partial^2 \mathbf{u}_\mathbf{i}}{\partial \mathbf{x}_\mathbf{j} \partial \mathbf{x}_\mathbf{k}}\right)\bigg|_{\mathbf{x}=\mathbf{x}^0}. \tag{8}$$

The microstress, $\mathbf{p}_{\boldsymbol{\alpha}}$, in DM is related to the microstrain, ε_α, as

$$\mathbf{p}_\alpha = \sum_{\beta=1}^{n} \mathbf{A}_{\alpha\beta} \varepsilon_\beta, \tag{9}$$

where $\mathbf{A}_{\alpha\beta}$ are the microlevel elastic moduli and n is the number of the neighboring nodes in the fundamental doublet bundle (for the configuration shown in Fig. 1, $n=6$). For the fundamental doublet bundle chosen above, $\mathbf{A}_{\alpha\beta}$ can be described by a 6×6 matrix $\mathbf{A}$ as shown below:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{12} & \mathbf{0} & \mathbf{A}_{15} & \mathbf{A}_{15} \\ \mathbf{A}_{12} & \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{15} & \mathbf{0} & \mathbf{A}_{15} \\ \mathbf{A}_{12} & \mathbf{A}_{12} & \mathbf{A}_{11} & \mathbf{A}_{15} & \mathbf{A}_{15} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{15} & \mathbf{A}_{15} & \mathbf{A}_{14} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{15} & \mathbf{0} & \mathbf{A}_{15} & \mathbf{0} & \mathbf{A}_{44} & \mathbf{0} \\ \mathbf{A}_{15} & \mathbf{A}_{15} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{44} \end{pmatrix}, \tag{10}$$

where

$$\mathbf{A}_{11} = \lambda + 4\mu \tag{11}$$

$$\mathbf{A}_{12} = \lambda + \mu, \tag{12}$$

$$\mathbf{A}_{44} = 4\mu, \tag{13}$$

$$\mathbf{A}_{15} = -2\mu, \tag{14}$$

and λ and μ are Lame's constants of the corresponding continuum mechanics. This particular choice of $\mathbf{A}_{\alpha\beta}$ yields an isotropic continuative relation at the continuous scale when $M=1$ (Ferrari, 2000). The macrostress $\sigma^{(M)}_{\mathbf{k}_1\mathbf{i}}$ is related to microstress $\mathbf{p}_{\boldsymbol{\alpha}\mathbf{i}}$ by the following equation (Ferrari *et al.*, 1997):

$$\sigma^{(M)}_{\mathbf{k}_1\mathbf{i}} = \sum_{\alpha=1}^{n} \tau^0_{\alpha\mathbf{k}_1} \sum_{\chi=1}^{M} (-1)^{\chi+1} \times \left[\frac{(\eta_\alpha)^{\chi-1}}{\chi!} \tau^0_{\alpha\mathbf{k}_2} \cdots \tau^0_{\alpha\mathbf{k}\chi} \frac{\partial^{\chi-1} \mathbf{p}_{\alpha\mathbf{i}}}{\partial \mathbf{x}_{\mathbf{k}_2} \cdots \partial \mathbf{x}_{\mathbf{k}_\chi}}\right]. \tag{15}$$

Using $\mathbf{p}_{\alpha\mathbf{i}} = \mathbf{p}_\alpha \tau^0_{\alpha\mathbf{i}}$, Eq. (15) leads to

$$\sigma_{\mathbf{ij}} = \sum_{\alpha=1}^{n} \tau^0_{\alpha\mathbf{i}} \tau^0_{\alpha\mathbf{j}} \mathbf{p}_\alpha, \tag{16}$$

for $M=1$ and

$$\sigma_{\mathbf{ij}} = \sum_{\alpha=1}^{n} \tau^0_{\alpha\mathbf{i}} \left(\tau^0_{\alpha\mathbf{j}} \mathbf{p}_\alpha - \frac{1}{2} \eta_\alpha \tau^0_{\alpha\mathbf{j}} \tau^0_{\alpha\mathbf{k}} \frac{\partial \mathbf{p}_\alpha}{\partial \mathbf{x}_\mathbf{k}}\right), \tag{17}$$

for $M=2$.

The conservation of the linear momentum (Ferrari, 1997) leads to the following wave equation for DM:

$$\sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \sum_{\chi=1}^{M} (-1)^{\chi-1} \mathbf{A}_{\alpha\beta} \tau^0_{\alpha\mathbf{i}} \frac{(\eta_\alpha)^{\chi-1}}{\chi!} \tau^0_{\alpha\mathbf{k}_1} \cdots \tau^0_{\alpha\mathbf{k}_\chi} \times \frac{\partial^\chi \varepsilon_\beta}{\partial \mathbf{x}_{\mathbf{k}_1} \cdots \partial \mathbf{x}_{\mathbf{k}_\chi}} = \rho \frac{\partial^2 \mathbf{u}_\mathbf{i}}{\partial \mathbf{t}^2}, \tag{18}$$

where ε_β is the microstrain given by Eq. (5) and ρ is the mass density of the material. The wave equation (18) can be reduced to $M=1$ [Eq. (19)] and $M=2$ [Eq. (21)], respectively, as

$$\mathbf{C}_{\mathbf{ijkl}} \frac{\partial^2 \mathbf{u}_\mathbf{j}}{\partial \mathbf{x}_\mathbf{k} \partial \mathbf{x}_\mathbf{i}} = \rho \frac{\partial^2 \mathbf{u}_\mathbf{i}}{\partial \mathbf{t}^2}, \tag{19}$$

where

$$\mathbf{C}_{\mathbf{ijkl}} = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \mathbf{A}_{\alpha\beta} \tau^0_{\alpha i} \tau^0_{\alpha j} \tau^0_{\beta k} \tau^0_{\beta l}, \tag{20}$$

and

$$\mathbf{C}_{\mathbf{ijkl}} \frac{\partial^2 \mathbf{u}_\mathbf{j}}{\partial \mathbf{x}_\mathbf{k} \partial \mathbf{x}_\mathbf{l}} - \mathbf{C}_{\mathbf{ijklmn}} \frac{\partial^4 \mathbf{u}_\mathbf{j}}{\partial \mathbf{x}_\mathbf{k} \partial \mathbf{x}_\mathbf{l} \partial \mathbf{x}_\mathbf{m} \partial \mathbf{x}_\mathbf{n}} = \rho \frac{\partial^2 \mathbf{u}_\mathbf{i}}{\partial \mathbf{t}^2}, \tag{21}$$

where

$$\mathbf{C}_{\mathbf{ijklmn}} = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \mathbf{A}_{\alpha\beta} \frac{\eta_\alpha^2}{4} \tau^0_{\alpha i} \tau^0_{\alpha j} \tau^0_{\alpha k} \tau^0_{\beta l} \tau^0_{\beta m} \tau^0_{\beta n}. \tag{22}$$

1. Dispersion of longitudinal waves

Let us assume a longitudinal plane wave propagating in the $\mathbf{x}_1$ direction, i.e.

$$\mathbf{u}_1 = \mathbf{u}_{10} \exp[\mathbf{i}(\omega \mathbf{t} - \mathbf{k}\mathbf{x}_1)], \tag{23}$$

where ω and k are angular frequency and wave number, respectively. Substituting Eq. (23) into Eq. (18), we obtain

$$\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\sum_{S=1}^{M}\sum_{\chi=1}^{M}(-1)^{\chi}(\mathbf{i})^{\chi+S}\mathbf{A}_{\alpha\beta}(\tau_{\alpha1}^{0})^{\chi+1}(\tau_{\beta1}^{0})^{S+1}\times\frac{(\eta_{\alpha})^{\chi-1}(\eta_{\beta})^{S-1}}{\chi!\mathbf{S}!}\mathbf{k}^{\chi+S}=\rho\omega^{2}. \tag{24}$$

Noting that the Mth-order longitudinal wave phase velocity $C_{L,M}$ is equal to ω/k, from Eq. (24), we obtain

$$\mathbf{C}_{\mathbf{L,M}}=\left[\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\sum_{S=1}^{M}\sum_{\chi=1}^{M}(-1)^{\chi}(\mathbf{i})^{\chi+S}\times\mathbf{A}_{\alpha\beta}(\tau_{\alpha1}^{0})^{\chi+1}(\tau_{\beta1}^{0})^{S+1}\times\frac{(\eta_{\alpha}/\lambda)^{\chi-1}(\eta_{\beta}/\lambda)^{S-1}(2\pi)^{\chi+S-2}}{\rho\chi!\mathbf{S}!}\right]^{1/2}. \tag{25}$$

For $M=1$, from Eq. (25) we can get

$$\mathbf{C}_{L,1}=\left[\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\mathbf{A}_{\alpha\beta}\frac{(\tau_{\alpha1}^{0})^{2}(\tau_{\beta1}^{0})^{2}}{\rho}\right]^{1/2}=\sqrt{\frac{\mathbf{A}_{11}-\mathbf{A}_{44}/2}{\rho}}=\sqrt{\frac{\lambda+2\mu}{\rho}}, \tag{26}$$

which is in agreement with the longitudinal phase velocity in CCM. For $M=2$, from Eq. (25), we obtain

$$\mathbf{C}_{L,2}=\sqrt{\frac{\mathbf{A}_{11}-\mathbf{A}_{44}/2}{\rho}+\frac{1}{4\rho}\left(\frac{\eta_{\alpha}}{\lambda}\right)^{2}\left(\mathbf{A}_{11}-\frac{\mathbf{A}_{44}}{\sqrt{2}}+\frac{\mathbf{A}_{44}}{4}\right)(2\pi)^{2}}=\mathbf{C}_{L,1}\sqrt{1+\frac{\delta}{\lambda^{2}}}, \tag{27}$$

$$\delta=\pi^{2}\eta_{\alpha}^{2}\left(\mathbf{A}_{11}-\frac{\mathbf{A}_{44}}{\sqrt{2}}+\frac{\mathbf{A}_{44}}{4}\right)\Big/(\mathbf{A}_{11}-\mathbf{A}_{44}/2). \tag{28}$$

It is clear that $\mathbf{C}_{L,2}$ is a function of wavelength. Since $\mathbf{A}_{11}=\lambda+4\boldsymbol{\mu}$ and $\mathbf{A}_{44}=4\boldsymbol{\mu}$, $\delta>0$, thus $\mathbf{C}_{L,2}$ is an increasing function of frequency; i.e., the dispersion shows an upswing relation.

From Eq. (25) and assuming $\eta_{\alpha}=\eta_{\beta}=\eta$ and $M=\infty$, we can obtain

$$\mathbf{C}_{\mathbf{L},\infty}=\sqrt{\mathbf{A}_{11}/\rho}\frac{\sin(\pi\eta/\lambda_{0})}{(\pi\eta/\lambda_{0})}. \tag{29}$$

This agrees with the dispersion relation with periodicity $k=2\pi/\eta$ for a one-dimensional longitudinal wave propagating in a monatomic lattice of spacing η (Kittel, 1971). It confirms that for $M=\infty$, DM is equivalent to Lattice Mechanics for a periodic and discrete crystal.

2. Dispersion of shear waves

If a shear plane wave propagating in the $\mathbf{x}_1$ direction and the displacement is along the x_2 direction,

$$\mathbf{u}_2=\mathbf{u}_{20}\exp[\mathbf{i}(\omega\mathbf{t}-\mathbf{k}\mathbf{x}_1)]. \tag{30}$$

Substituting Eq. (30) into Eq. (18), we obtain

$$\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\sum_{\mathbf{S}=1}^{M}\sum_{\chi=1}^{M}(-1)^{\chi}(\mathbf{i})^{\chi+\mathbf{S}}\mathbf{A}_{\alpha\beta}\tau_{\alpha2}^{0}(\tau_{\alpha1}^{0})^{\chi}\tau_{\beta2}^{0}(\tau_{\beta1}^{0})^{\mathbf{S}}\times\frac{(\eta_{\alpha})^{\chi-1}(\eta_{\beta})^{\mathbf{S}-1}}{\chi!\mathbf{S}!}\mathbf{k}^{\chi+\mathbf{S}}=\rho\omega^{2}. \tag{31}$$

Once again, noting that the shear wave phase velocity $C_{S,M}$ is equal to ω/k, from Eq. (31), we get

$$\mathbf{C}_{\mathbf{S,M}}=\left[\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\sum_{\mathbf{S}=1}^{M}\sum_{\chi=1}^{M}(-1)^{\chi}(\mathbf{i})^{\chi+\mathbf{S}}\times\mathbf{A}_{\alpha\beta}\tau_{\alpha2}^{0}(\tau_{\alpha1}^{0})^{\chi}\tau_{\beta2}^{0}(\tau_{\beta1}^{0})^{\mathbf{S}}\times\frac{(\eta_{\alpha}/\lambda)^{\chi-1}(\eta_{\beta}/\lambda)^{\mathbf{S}-1}(2\pi)^{\chi+\mathbf{S}-2}}{\rho\chi!\mathbf{S}!}\right]^{1/2}. \tag{32}$$

For $M=1$, from Eq. (32) we can obtain

$$\mathbf{C}_{\mathbf{S},1}=\left[\frac{1}{\rho}\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\mathbf{A}_{\alpha\beta}\tau_{\alpha1}^{0}\tau_{\alpha2}^{0}\tau_{\beta1}^{0}\tau_{\beta2}^{0}\right]^{1/2}=\sqrt{\frac{\mathbf{A}_{44}/4}{\rho}}=\sqrt{\frac{\mu}{\rho}}. \tag{33}$$

Equation (33) is the same equation used for the shear wave phase velocity in CCM. For $M=2$, we get

$$\mathbf{C}_{\mathbf{S},2}=\sqrt{\frac{\mathbf{A}_{44}/4}{\rho}+\frac{1}{\rho}\left(\frac{\eta_{\alpha}}{\lambda}\right)^{2}\frac{\mathbf{A}_{44}}{32}(2\pi)^{2}}. \tag{34}$$

B. The dispersion-attenuation model

According to an equation developed from a Kramers–Kronig relation (Szabo and Wu, 2000), if attenuation associated with a medium is a function of frequency, f, and can be fitted into a power law, i.e.,

$$\alpha(\mathbf{f})=\alpha_{0}+\alpha_{1}\mathbf{f}^{y}, \tag{35}$$

and y is an even integer or noninteger, attenuation coefficient, α_1, and dispersion (excess of phase velocity), $\delta c(f)$, are related by the following equations:

$$\delta\mathbf{c}(\mathbf{f})=\mathbf{c}(\mathbf{f})-\mathbf{c}(\mathbf{f}_{0})=-\frac{[\mathbf{c}(\mathbf{f}_{0})]^{2}\alpha_{1}\tan(\pi\mathbf{y}/2)}{2\pi*8.686*10^{4}}(|\mathbf{f}|^{\mathbf{y}-1}-|\mathbf{f}_{0}|^{\mathbf{y}-1}), \tag{36}$$

where $\mathbf{f}_0$ is a reference frequency and the units for α_1, $c(\mathbf{f})$ and $\mathbf{f}$ are dB/(cm MHz), m/s, and MHz, respectively.

II. COMPUTER SIMULATION

Soft tissue and polymers are discrete and granular on a cellular and a molecular scale, respectively; consequently, they should be good candidates for DM (Liu and Ferrari, 2003). The above mathematical framework of DM was used to describe them. The following parameters were used in our numerical calculations: (1) For fresh human brain tissue, $\rho=1012$ kg/m^3, $A_{11}=\lambda+4\mu=2.72$ GPa, and $A_{44}=4\mu=0.5$ GPa. (2) For low-density polyethylene, $\rho=900$ kg/m^3, $A_{11}=\lambda+4\mu=8.83$ GPa, and $A_{44}=4\mu=5.832$ GPa. Elastic constants A_{11} and A_{44} for fresh brain

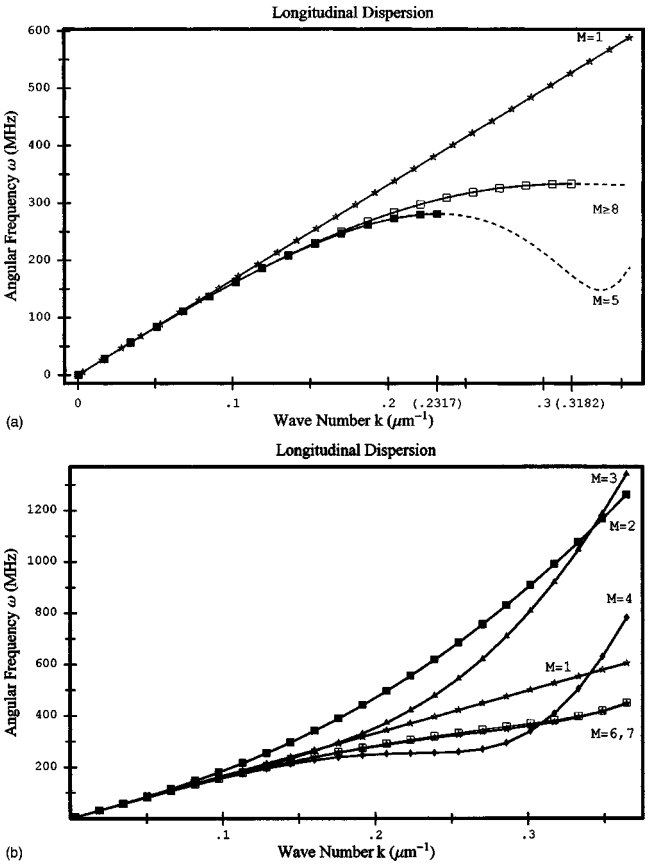


FIG. 3. Dispersion relation curves of longitudinal waves: (a) for $M=1$, 5, and $\geqslant 8$. When $M \geqslant 8$, they fall one on top of the other and become indistinguishable. For $M=5$ and $\geqslant 8$, the van Hove singularities occur at $k=2.317/\eta$, $3.182/\eta$, respectively. If the wave number $k>2.317/\eta$ (or $3.182/\eta$), the dispersion becomes unphysical. Therefore, curves of $M=5$, and $\geqslant 8$ stop at when $k=2.317/\eta$ (the curve becomes a broken-line curve beyond $2.317/\eta$) and $3.182/\eta$, respectively. (b) For $M=1$, 2, 3, 4, 6, and 7. No van Hove singularity for each curve.

tissue were measured by Kremkau *et al.* (1981). Elastic constants for low density polyethylene were measured by Wu (1996). For mathematical simplicity, we also assumed that the internodal distances, η_α, are all equal, i.e., $\eta_\alpha=\eta$ for $\alpha=1,...,6$ for all calculations below. For brain tissue, $\eta=10\ \mu$m and for low-density polyethylene $\eta=0.3\ \mu$m (Tobolsky, 1960).

A. Longitudinal waves

Figure 3 contains two plots that illustrate the dispersion relation (ω vs k) of the longitudinal waves for materials that have the fundamental doublet bundle discussed above and also $\eta=10\ \mu$m. Plots in Fig. 3(a) include the dispersion relations for $M=1$ (the classical continuum case), $M=5$, and $M\geqslant 8$ and plots in Fig. 3(b) include those for $M=2,3,4,6,7$ (the same $M=1$ case was included again for reference). It is well known that classical continuum mechanics predicts that bulk ultrasound waves are nondispersive, i.e., ω vs k is a straight line. It should be noted that for the longitudinal waves only the dispersion relations shown in Fig. 3(a), except $M=1$, have so-called van Hove singularities (Kittel, 1971), i.e., $\partial\omega/\partial\mathbf{k}=0$ (the group velocity equals zero). For $M=8$, the van Hove singularity occurs at $\mathbf{k}=3.182/\eta$, which is near the boundary ($\mathbf{k}=\pi/\eta$) of the first Brillouin zone of

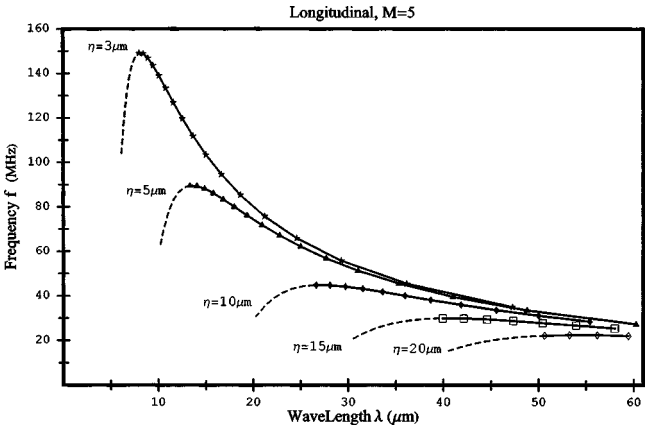


FIG. 4. Frequency versus wavelength for various η. The location where $\partial\mathbf{f}/\partial\lambda=0$ sets the highest frequency for longitudinal waves to propagate in the discrete material. The dashed line parts of the curves are unphysical.

a monatomic lattice in solid state physics (Kittel, 1971). Just like Lattice Mechanics, it restricts the wavelength λ by $\eta k_{\max}\leqslant 3.182$, i.e., $\lambda\geqslant 1.974\eta$ for $M=8$. For $M=5$, the van Hove singularity occurs at $\mathbf{k}=2.317/\eta$; thus, the restriction now is $\lambda\geqslant 2.712\eta$.

Dispersion relations may be plotted in a different way. Figure 4 shows the relation between f and λ for various η when $M=5$. It is evident that similar van Hove singularities occur at different λ for various η. It also clearly shows that the van Hove singularity sets the highest frequency limit, $f_{\max}$, or the lowest wavelength, $\lambda_{\min}$, for a known η. Above $f_{\max}$ or below $\lambda_{\min}$ the dispersion becomes unphysical (dashed line parts of curves, two wavelengths correspond to one frequency), i.e., a plane longitudinal wave cannot propagate in this material. Figure 4 also show that for $\eta=10\ \mu$m, $f_{\max}=44.76$ MHz, which is very accessible in practical ultrasound applications. Values of $\eta k_{\max}$ and $f_{\max}$ at van Hove singularities for various M are listed in Table I; the values of $\eta k_{\max}$ and $f_{\max}$ are calculated for $\eta=10\ \mu$m.

Figure 5 contains curves of the normalized longitudinal wave phase velocities, $C_{L,M}/C_{L,1}$ ($C_{L,1}=\sqrt{\lambda+2\mu/\rho}=1560$ m/s for this material), as a function of η/λ for different M, where $C_{L,M}$ ($M=1,2,...$) is the phase velocity of Mth order given by Eq. (25). It is evident from Fig. 5 that for $M=1$, the longitudinal wave phase, $C_{L,1}$, is independent of η/λ i.e., it is nondispersive and its value is given by Eq. (26). The longitudinal wave phase velocities for $M=5$, 6, 7 and $M\geqslant 8$ decrease as η/λ increases initially if $\eta/\lambda<0.55$ but they all become smaller than $C_{L,1}$ in the range of η/λ shown in the figure (for $M=6$ and 7, they eventually become

TABLE I. Values of wave number and frequency at the van Hove singularity, for longitudinal waves. $f_{\max}$ are in MHz and are evaluated at $\eta=10\ \mu$m.

M	$(\eta)\ k_{\max}$	$f_{\max}$	M	$(\eta)\ k_{\max}$	$f_{\max}$
1	⋯	⋯	7	⋯	⋯
2	⋯	⋯	8	3.182	53.11
3	⋯	⋯	9	3.022	52.98
4	⋯	⋯	10	3.206	53.65
5	2.317	44.76	25	3.203	53.61
6	⋯	⋯	∞	3.203	53.61

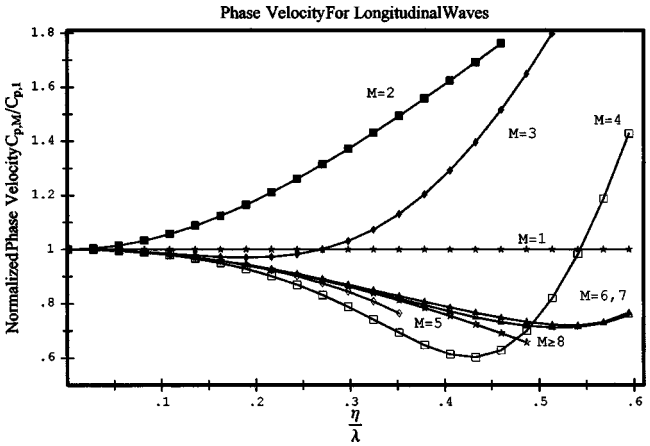


FIG. 5. The normalized phase velocity $C_{L,M}/C_{L,1}$, vs η/λ for longitudinal waves.

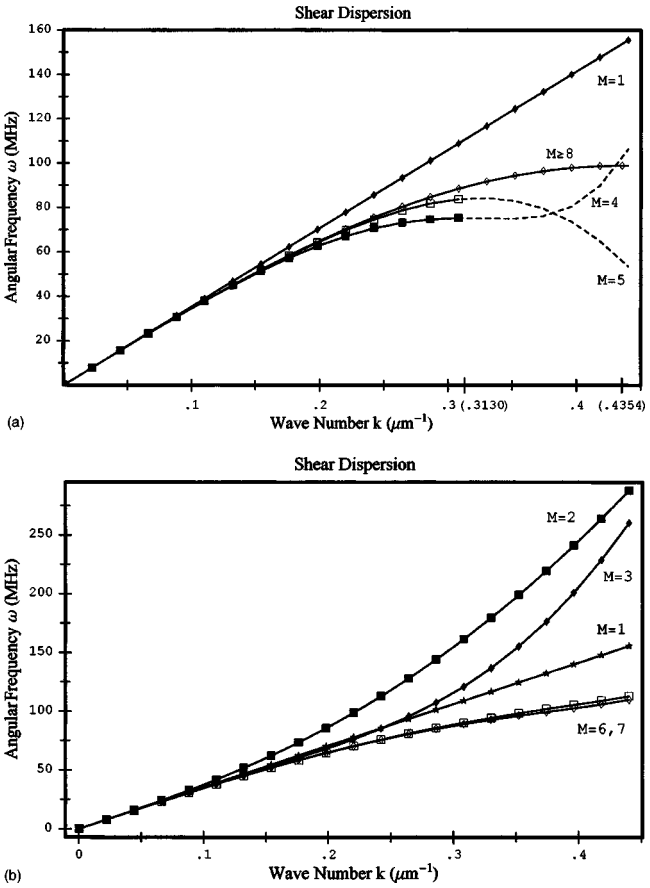


FIG. 6. Dispersion relation curves of shear waves: (a) for $M=1$, 4, 5, and $\geqslant 8$. When $M \geqslant 8$, they fall one on top of the other and become indistinguishable. For $M=4$, 5, and $\geqslant 8$, van Hove singularities occur at $k=3.130/\eta$, $3.262/\eta$, and $4.354/\eta$, respectively. If wave number $k>3.130/\eta$ (or $3.262/\eta$ or $4.354/\eta$,) the dispersion becomes unphysical. Therefore, curves of $M=4$, 5, and $\geqslant 8$ stop at $3.130/\eta$, $3.262/\eta$ or $4.354/\eta$, respectively. (b) For $M=1$, 2, 3, 6, and 7. No van Hove singularity for those curves.

greater than $C_{L,1}$ if $\eta/\lambda \gg 0.6$). For $M=2$, the phase velocity is an increasing function of η/λ and it is always above $C_{L,1}$. For $M=3$, the phase velocity $C_{L,3}$ is smaller than $C_{L,1}$ and becomes larger than $C_{L,1}$ when $\eta/\lambda > 0.28$. For $M=4$, 6, and 7 the shapes of the dispersion curves are very similar to that of $M=3$; the difference in $C_{L,4}$ becomes larger than $C_{L,1}$ when $\eta/\lambda > 0.54$; and $C_{L,6}$ and $C_{L,7}$ become larger than $C_{L,1}$ when η/λ reaches even higher values (the out of scale range of the figure). It should be noted that the dispersion curves in Fig. 5 terminate at their respective $\eta/\lambda_{\min}$ coresponding to their van Hove singularities for $M=5$ and $M \geqslant 8$. There is no limitation of η/λ for the dispersion of other values of M.

B. Shear waves

Figure 6 illustrates the dispersion relation (ω vs k) of shear waves for the same materials ($\eta = 10\ \mu$m). Plots in Fig. 6(a) include the dispersion relation for $M=1$ (the classical continuum case), $M=4$, 5 and $M \geqslant 8$, and plots in Fig. 6(b) include those for $M=2$, 3, 6, 7 (the same $M=1$ case was also included). Once again, the result obtained from Eq. (33) for $M=1$, as shown in Fig. 6, indicates that the system is nondispersive for $M=1$, which is in agreement with the description of shear waves in classical continuum mechanics. Curves in Fig. 6 show that the system becomes dispersive for $M>1$. Only for $M=4$, 5 and $M \geqslant 8$ shown in Fig. 6(a), van Hove singularities occur; they occur at $\mathbf{k}=3.130/\eta$, $3.262/\eta$, and $4.443/\eta$, respectively. Values of $(\eta)k_{\max}$ and $f_{\max}$ at van Hove singularities for various M are listed in Table II; the values of $(\eta)k_{\max}$ and $f_{\max}$ are also calculated for $\eta = 10\ \mu$m.

Figure 7 shows the relation between f and λ for various η when $M=5$. Just like its longitudinal wave counterpart, it is evident that similar van Hove singularities occur at different λ for various η. It also clearly shows that the van Hove singularity sets the highest frequency limit, $f_{\max}$, or the lowest wavelength, $\lambda_{\min}$, for a known η. Above $f_{\max}$ or below $\lambda_{\min}$ the dispersion becomes unphysical, i.e., a plane shear wave cannot propagate in this material. Figure 7 also shows that for $\eta = 10\ \mu$m, $f_{\max}=13.41$ MHz, which is easily accessible in practical ultrasound applications.

Figure 8 contains curves of the normalized shear wave phase velocities, $C_{S,M}/C_{S,1}$ ($C_{S,1}=\sqrt{\mu/\rho}=351$ m/s for this material) as a function of η/λ for different M. It is evident, just like the longitudinal wave for $M=1$, that the shear wave phase velocity, $C_{S,1}$, is independent of η/λ i.e., it is nondispersive. The shear wave phase velocities of all M except $M=2$, 3 are smaller than that of $M=1$, $C_{S,1}$, if $\eta/\lambda < 0.8$ (for $M=6$ and 7, they eventually become greater than $C_{S,1}$ if $\eta/\lambda \gg 0.8$). For $M=3$, the shear wave phase velocity starts below $C_{S,1}$ and becomes larger than $C_{S,1}$ when $\eta/\lambda > 0.38$. It should be noted that the curves terminate at their respective $\eta/\lambda_{\max}$ corresponding to their van Hove singularities for M

TABLE II. Values of wave number and frequency at the van Hove singularity, for shear waves. $f_{\max}$ are in MHz and are evaluated at $\eta = 10\ \mu$m.

M	$(\eta)\ k_{\max}$	$f_{\max}$	M	$(\eta)\ k_{\max}$	$f_{\max}$
1	⋯	⋯	7	⋯	⋯
2	⋯	⋯	8	4.354	15.74
3	⋯	⋯	9	4.284	15.76
4	3.130	11.98	10	4.453	15.93
5	3.262	13.41	25	4.443	15.92
6	⋯	⋯	∞	4.443	15.92

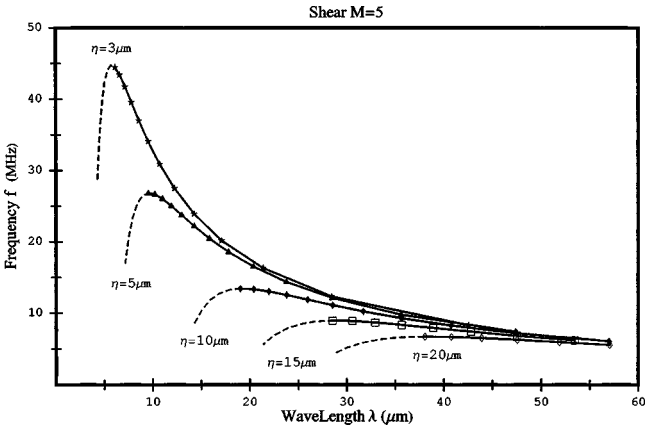


FIG. 7. Frequency versus wavelength for various η. The location where $\partial \mathbf{f}/\partial \lambda = 0$ sets the highest frequency for shear waves to propagate in the discrete material. The dashed line parts of the curves are unphysical.

$=4$, 5, and $M \geqslant 8$. There is no limitation of η/λ for other values of M.

C. Comparison between experimental data and results predicted by DM and attenuation models

Figure 9(a) is used to compare results calculated from the DM model [Eq. (25)] and the attenuation model [Eq. (36)] for fresh brain tissue ($\eta = 10\ \mu$m). Apparently, below 4 MHz, dispersion predicted by the attenuation model is greater than that calculated from the DM model. The DM results grow much faster when frequency is greater than 4 MHz. On the contrary, dispersion calculated using the attenuation model almost reaches a plateau after 4 MHz. We replotted the results in the frequency range between 1–5 MHz in a smaller velocity scale in Fig. 9(b). In Fig. 9(b), we also included the values measured by Kremkau *et al.* (1981). Limited by the number of measured data, a detailed comparison was not possible. Nevertheless, it can be argued from the figure that both theoretical models can explain data equally well within the experimental errors.

Similar plots were made for low-density polyethylene ($\eta = 0.3\ \mu$m) in Fig. 10(a). Dispersion for low-density polyethylene predicted by the attenuation model is greater than the DM model when 5 MHz$< f <$900 MHz. Dispersion predicted by DM becomes dominant when $f >$900 MHz. We also plotted dispersion for 2 MHz$< f <$8 MHz, as shown in Fig. 10(b). In Fig. 10(b), experimental data (Wu, 1996) were also included. Since the central frequency of a broadband pulse used in the experiment was 4.8 MHz, $f_0 = 4.8$ MHz was used as a reference frequency. From Fig. 10(b), it is evident that dispersion predicted by DM was negligible and results from the attenuation model agree well with the experimental data.

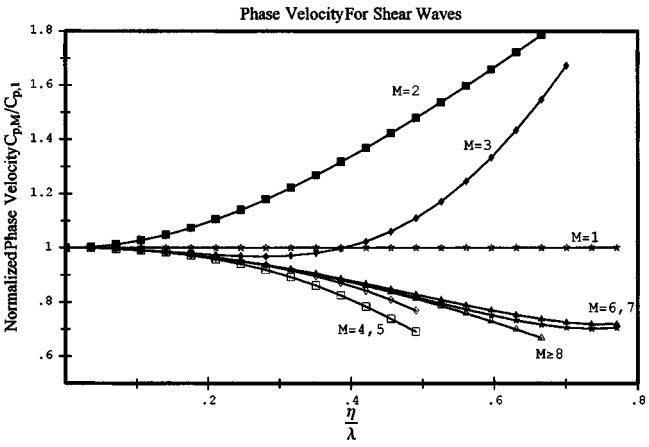


FIG. 8. The normalized phase velocity $C_{S,M}/C_{S,1}$, vs η/λ for shear waves.

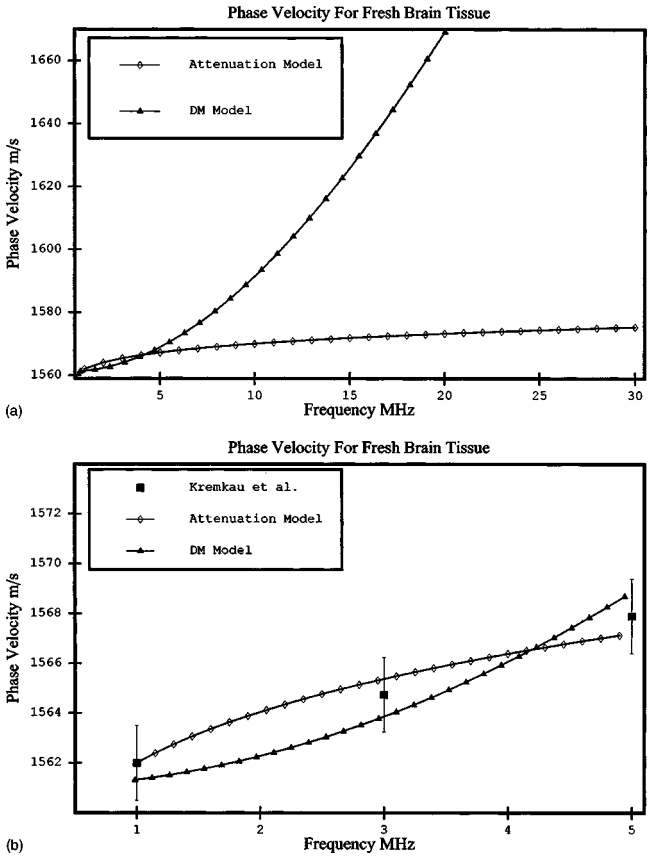


FIG. 9. Dispersion (excess phase velocity) as a function of frequency for fresh human brain tissue and $\eta = 10\ \mu$m. Experimental data: ■; DM: ▲; attenuation model: ◇. (a) 1 MHz$< f <$20 MHz. (b) 1 MHz$< f <$5 MHz.

III. DISCUSSION

A multiscale wave equation was derived from Doublet Mechanics. The dispersion relations for longitudinal waves and shear waves were obtained for a bundle structure commonly used to describe isotropic discrete doublets. The van Hove singularities for longitudinal and shear waves determined above are for this particular microbundle structure. As described earlier, the van Hove singularities set the highest frequency of a plane wave that can propagate in this material. The general rule seems to be that the smaller the internodal distance η is the higher is $f_{\max}$.

When DM is used to model a discrete or partially discrete system, an investigation of dispersion relations should be seriously considered. When longitudinal and shear wave dispersion relations of some biological soft tissue (Bamber, 1981; Kremkau *et al.*, 1981) and polymer materials (Wu, 1996; He, 1999) were measured at frequencies <10 MHz,

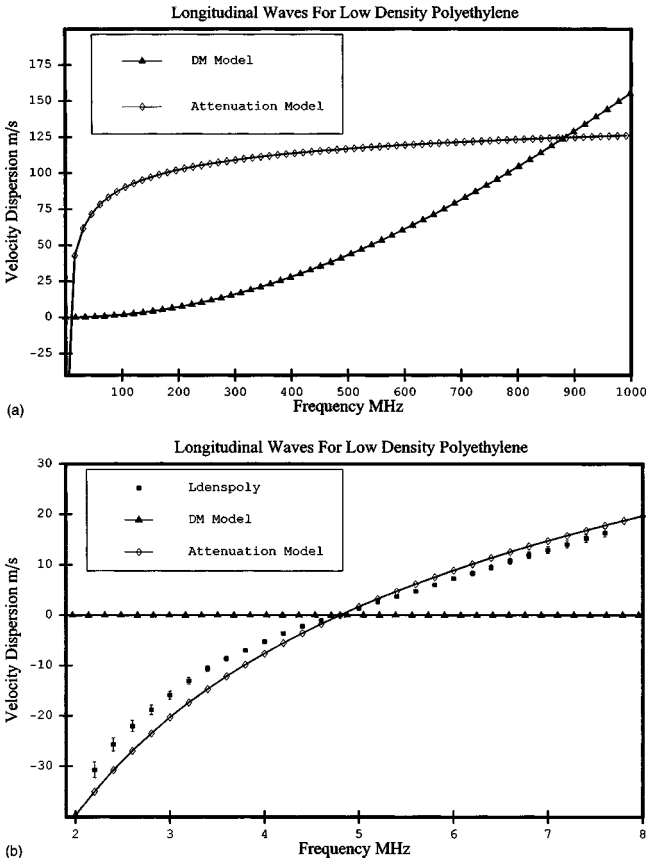


FIG. 10. Dispersion (excess phase velocity) as a function of frequency for low-density polyethylene and $\eta = 0.3\ \mu$m. Experimental data: ■; DM: ▲; attenuation model: ◇. (a) 5 MHz$< f <$100 MHz. (b) 2 MHz$< f <$8 MHz.

they all showed very similar behavior as those of $M = 2$ described above; i.e., the phase velocity is an increasing function of f or $(1/\lambda)$. Therefore, the DM theory of $M = 2$ may be appropriate to describe those biological tissue and polymers, which suggests that the structure of soft tissue and polymers lay between a perfect continuum and perfect crystal.

From Figs. 9 and 10, we have observed that the discreteness of a system may introduce dispersion; it is especially true at high frequencies. When the wavelength of ultrasound becomes close to the internodal distance η, dispersion associated with the discrete nature of a system may become important and dominant. On the other hand, if wavelength is much smaller than η(at low frequency), dispersion due to discreteness may be negligible. Due to the lack of experimental data in the literature, a comparison between theory and experiments at high frequencies was not possible. The validity of DM would be tested by such comparisons in the future.

It is also important to realize that to apply an adequate Mth-order DM to a discrete system, the dispersion relation of this material may need to be measured first.

Auld, B. A. (**1990**). *Acoustic Fields and Waves in Solids*, 2nd ed. (Krieger, Malabar, Fl), pp. 251–253.

Bamber, J. C. (**1986**). "Speed of sound," in *Physical Principles of Medical Ultrasonics*, edited by C. R. Hill (Ellis Horwood Limited, New York), pp. 200–224.

Ferrari, M., Granik, V. T., and Imam, A. (**1997**). "Introduction to doublet mechanics," in *Advances in Doublet Mechanics*, edited by M. Ferrari, V. T. Granik, and A. Imam (Springer-Verlag, Berlin), pp. 1–26.

Ferrari, M. (**2000**). "Nanomechanics, and biomedical nanomechanics: Eshelby's inclusion and inhomogeneity problems at the discrete/continuum interface," Biomed. Microdevices **2**, 273–281.

Ferrari, M. (**2003**). Private communication.

Granik, V. T. (**1997**). "Comparison with other theories," in *Advances in Doublet Mechanics*, edited by M. Ferrari, V. T. Granik, and A. Imam (Springer-Verlag, Berlin), pp. 51–82.

He, P. (**1999**). "Experimental verification of models for determining dispersion from attenuation," IEEE Trans. UFFC 46, 706–714.

He, P. (**2000**). "Measurement of acoustic dispersion using both transmitted and reflected pulses," J. Acoust. Soc. Am. **107**, 801–806.

Kittel, C. (**1971**). *Introduction to Solid State Physics* (Wiley, New York), p. 214.

Kremkau, F. W., Barnes, R. W., and McGraw, C. P. (**1981**). "Ultrasonic attenuation and propagation speed in normal human brain," J. Acoust. Soc. Am. **70**, 29–38.

Liu, J., and Ferrari, M. (**2002**). "Mechanical spectral signatures of malignant disease? A small-sample, comparative study of continuum vs. nano-biomechanical data analyses," Disease Markers **18**, 175–183.

O'Donnell, M., Jaynes, E. T., and Miller, J. G. (**1981**). "Kramers-Kronig relationship between ultrasonic attenuation and velocity," J. Acoust. Soc. Am. **69**, 696–701.

Szabo, T. L. (**1994**). "Time domain wave equations for lossy media obeying a frequency power law," J. Acoust. Soc. Am. **96**, 491–500.

Szabo, T. L., and Wu, J. (**2000**). "A model for longitudinal and shear wave propagation in viscoelastic media," J. Acoust. Soc. Am. **107**, 2437–2446.

Tobolsky, A. T. (**1960**). *Properties and Structure of Polymers* (Wiley, New York).

Wu, J. (**1996**). "Determination of velocity and attenuation of shear waves using ultrasonic spectroscopy," J. Acoust. Soc. Am. **99**, 2871–2875.

Echolocation signals of free-ranging killer whales (*Orcinus orca*) and modeling of foraging for chinook salmon (*Oncorhynchus tshawytscha*)

Whitlow W. L. Au
Marine Mammal Research Program, Hawaii Institute of Marine Biology, University of Hawaii, Kaneohe, Hawaii 96744

John K. B. Ford
Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, British Columbia, V9T 6N7 Canada

John K. Horne
School of Aquatic and Fisheries Science, University of Washington, Seattle, Washington

Kelly A. Newman Allman
The Marine Mammal Center, Sausalito, California 94965

(Received 3 April 2003; revised 7 November 2003; accepted 24 November 2003)

Fish-eating ''resident''-type killer whales (*Orcinus orca*) that frequent the coastal waters off northeastern Vancouver Island, Canada have a strong preference for chinook salmon (*Oncorhynchus tshawytscha*). The whales in this region often forage along steep cliffs that extend into the water, echolocating their prey. Echolocation signals of resident killer whales were measured with a four-hydrophone symmetrical star array and the signals were simultaneously digitized at a sample rate of 500 kHz using a lunch-box PC. A portable VCR recorded the images from an underwater camera located adjacent to the array center. Only signals emanating from close to the beam axis (1185 total) were chosen for a detailed analysis. Killer whales project very broadband echolocation signals (Q equal 0.9 to 1.4) that tend to have bimodal frequency structure. Ninety-seven percent of the signals had center frequencies between 45 and 80 kHz with bandwidths between 35 and 50 kHz. The peak-to-peak source level of the echolocation signals decreased as a function of the one-way transmission loss to the array. Source levels varied between 195 and 224 dB $re:1\ \mu$Pa. Using a model of target strength for chinook salmon, the echo levels from the echolocation signals are estimated for different horizontal ranges between a whale and a salmon. At a horizontal range of 100 m, the echo level should exceed an *Orcinus* hearing threshold at 50 kHz by over 29 dB and should be greater than sea state 4 noise by at least 9 dB. In moderately heavy rain conditions, the detection range will be reduced substantially and the echo level at a horizontal range of 40 m would be close to the level of the rain noise. [DOI: 10.1121/1.1642628]

PACS numbers: 43.80.Ev, 43.80.Ka, 43.80.Jz [FD] Pages: 901–909

I. INTRODUCTION

Echolocation experiments with captive dolphins have been conducted since the early 1950's (Schevill and Lawrence, 1953a,b), and yet very little is known about how dolphins utilize their unique echolocation capability in the wild. A variety of detection and discrimination experiments has been conducted (Au, 1993) using a wide assortment of targets including dead fish (Kellogg, 1958), metallic, foam, or plastic spheres, cylinders, and plates. (Au, 1993; Nachtigall and Moore, 1988). For example, an Atlantic bottlenose dolphin (*Tursiops truncates*) can detect a 7.62-cm diameter sphere out to 113 m, in a noisy bay dominated by snapping shrimp (Au and Snyder, 1980). In a quiet environment the detection range for the 7.62-cm sphere would be about 178 m (Au *et al.*, 2002). The echolocation capabilities of dolphins on these ''unnatural'' objects have provided a wealth of information. Yet we are left with the nagging issue of relevancy of these experiments in understanding how dolphins use echolocation to forage for prey.

One problem with studying echolocation in the wild is the difficulty of obtaining accurate measurements of echolocation signal by free-ranging dolphins. The echolocation beam pattern is relatively narrow and if signals are not measured close to the axis of the animal's beam, the signals will be distorted (Au, 1993). It is also extremely difficult to determine the range between a moving animal and the recording hydrophone. A range determination is essential in order to calculate the source level (sound-pressure level 1 m from the dolphin) of the signals.

These two problems have been successfully overcome by using a short-baseline array of four hydrophones arranged as a symmetrical star (Au and Herzing, 2003; Rasmussen *et al.*, 2002). Such an array geometry has been used successfully by Aubauer (1995) to track flying bats. With this array, a researcher can determine whether a received signal was measured along the major axis of a dolphin's transmission beam as well as estimate the range of the animal from the array.

Killer whales (*Orcinus orca*), found in the coastal waters of British Columbia, Washington State and southeastern

Alaska, are ideal candidates to investigate the details of echolocation signal production in a free-ranging odontocete. This region is inhabited by two distinct ecotypes of the species, each with a different dietary specialization (Ford *et al.*, 2000). The ''transient'' ecotype selectively hunts marine mammal prey, while the ''resident'' ecotype feeds on fishes, particularly salmon. Of the six species of salmon observed to be taken, chinook (*Oncorhynchus tshawytscha*) was the most common, presumably because of their large size and rich lipid content (Ford *et al.*, 1998). Acoustic studies by Barrett-Lennard (1996) have shown that transient killer whales seldom use echolocation signals, likely as a hunting strategy, but resident killer whales frequently emit echolocation clicks while foraging for salmon. In this study, results from echolocation signal measurements of free-ranging resident-type killer whales are presented and these results are incorporated into a model involving *Orcinus* foraging on chinook salmon. This study may be the first to model foraging by echolocating odontocetes.

II. PROCEDURE

A. Acoustic measurement system

A four-hydrophone array with the hydrophones arranged as a symmetrical star was used to measure echolocation signals by killer whales. The array structure resembled the letter ''Y,'' with each arm being 45.7 cm in length and separated by an angle of 120°. The arms of the array are constructed out of 1.27-cm-o.d. polyvinyl chloride (PVC) pipe with a spherical hydrophone connected to the end of each pipe and the cable running through the center of the pipe. Another hydrophone is connected at the geometric center of the ''Y.'' An underwater housing attached to the back of the hydrophone mounting plate contained an amplifier and line driver for each of the hydrophones. A multiconductor cable consisting of five coaxial and two dc power lines connected the array electronics to an adjustable battery operated amplifier-filter box that was housed on a boat. The echolocation signals were acquired using a four-channel analog-to-digital converter controlled by a transportable ''lunch-box'' PC housed on the boat. A CCD video camera in another underwater housing was mounted next to the center hydrophone. Additional details of the measurement system can be found in Au and Herzing (2003).

Data acquisition was achieved using two Gage-1210, 12-bit dual simultaneous sampling data acquisition boards housed in the lunch-box PC via two EISA slots. The data acquisition system sampled at 500 kHz with a pretrigger capability. When the computer signaled the Gage-1210 to collect data, four channels of acoustic signals were simultaneously and continuously digitized. Data channels were logged into separate circular memories on each Gage-1210 board. When an echolocation signal was detected by the center hydrophone, it triggered the data acquisition board. Two hundred pretriggered points and 200 post-trigger points were collected for each channel and downloaded to the computer. A total of 80 clicks could be downloaded for each episode before the data had to be stored in a file on the hard drive. A specially constructed ISA board was also used to measure the time interval between clicks being acquired and to trigger a light-emitting diode to indicate that clicks were being captured and recorded. The interclick interval data were also downloaded and stored on the hard drive. The time of capture (to the closest 18-ms interval of the computer timing system) of each click was also saved and stored on the hard drive. The clock on a portable VCR was synchronized to the computer's clock so that the video images were synchronized with the acoustic data.

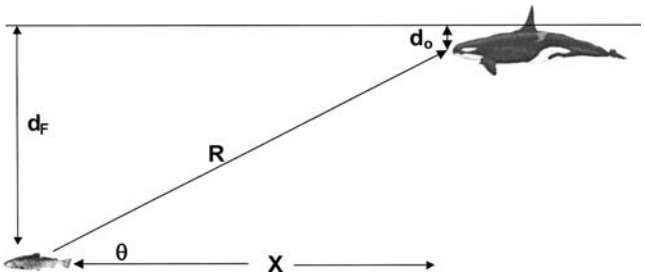


FIG. 1. Foraging geometry showing a killer whale at depth d_o in pursuit of a chinook salmon at range R and a depth d_F.

B. Acoustic measurements

Killer whale echolocation signals were measured on 4 consecutive days in Johnstone Strait, off northeastern Vancouver Island, Canada (50°32′N, 126°42′W) during August, 2000. The four-hydrophone array was deployed from a small boat while small groups of two to three killer whales foraged along the shoreline where rocky cliffs intersect the water at a relatively steep angle. When a group of whales was spotted, the boat would be positioned 150–200 m ahead of and in the path of the animals. The hydrophone array was lowered into the water with the center hydrophone at a depth between 1.2 and 1.5 m. At the same time, an operator controlled the data acquisition sequence by arming the computer to start data acquisition and also started the video tape recorder. As the whales approached the boat, they rarely deviated from their path except to veer to one side if the array was in their path so as to not strike the array.

C. Foraging geometry

In order to examine the use of echolocation by *Orcinus* while foraging on chinook salmon, the geometry in Fig. 1 will be used. In this geometry, the whale is in a pursuit mode directly behind the salmon. The following notations are used in Fig. 1: $\mathbf{d_F}$ and $\mathbf{d_o}$ are the depth of the salmon and the killer whale's sonar, respectively. $\mathbf{R}$ is the slant range between the fish and the *orca*, $\mathbf{X}$ is the horizontal range, and θ is the angle between the slant and horizontal range. The angle θ can be determined by the equation

$$\theta = \sin^{-1}\left(\frac{d_F - d_O}{R}\right) = \tan^{-1}\left(\frac{d_F - d_O}{x}\right). \quad (1)$$

The target strength of the salmon is a function of the angle θ, the frequency of the acoustic signal, and the length of the fish (Clay and Horne, 1994). Chinook salmon in Johnstone Strait typically swim at a mean depth of approximately 25 to 64 m during the day (Candy and Quinn, 1999). In this study

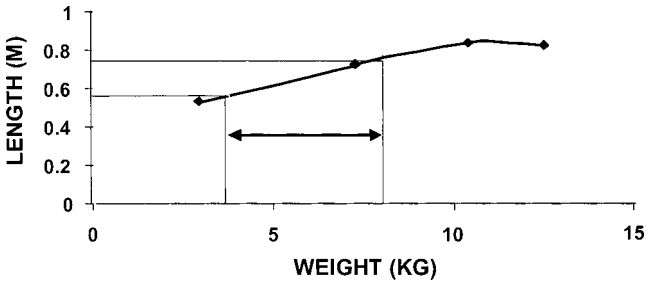


FIG. 2. Comparison of weight versus length for chinook salmon from data provided by Dr. Donald Rogers of the University of Washington.

we assume that the whale is close to the surface during its sonar search, so that a value of 1 m will be used for d_o in Eq. (1).

D. Calculation of salmon target strength

Data collected over a 20-year period in the nearshore waters between Vancouver Island and the mainland British Columbia, and around Queen Charlotte Islands, indicate that chinook salmon are the preferred prey of resident killer whales, during summer (Ford *et al.*, 1998). They also estimated that the preferred weight of the chinook salmon prey was between 3.7 and 8.1 kg. Using length and weight data (unpublished data courtesy of Donald Rogers, U.W., 2002) shown in Fig. 2, the corresponding length of a 3.7-kg chinook salmon is approximately 0.6 m, and 0.8 m for 8.1 kg. A length of 0.78 m will be used in this study.

In order to calculate the strength of an echo returning to an echolocating killer whale, the target strength of chinook salmon must be determined. Clay and Horne (1994) developed a Kirchhoff-ray mode (KRM) backscatter model in which the body of a fish is modeled as a series of short-length fluid-filled cylinders surrounding a set of gas-filled cylinders representing the swimbladder. Backscatter from each cylinder in the body and the swimbladder is estimated and then added coherently to estimate total backscatter as a function of fish length, orientation (i.e., tilt, roll), and acoustic carrier frequency. Model predictions have been shown to match backscatter measurements (Horne and Jech, 1999; Jech and Horne, 2002; Jech *et al.*, 1995). Target strengths of a chinook salmon at 50 kHz were calculated using digitized dorsal and lateral radiographs (Fig. 3) as input to a KRM model parametrized for a 0.78-m caudal length chinook salmon.

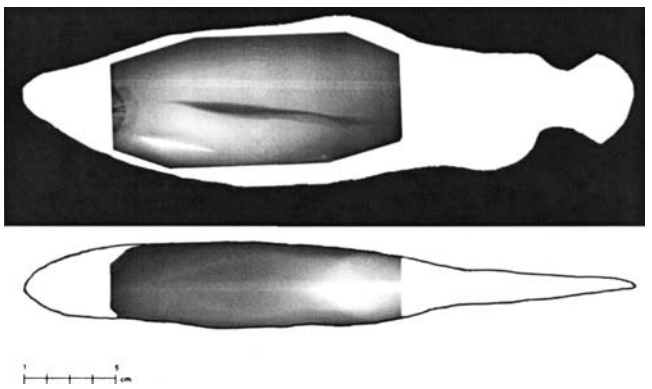


FIG. 3. Radiograph (x-ray images) of a chinook salmon. Top panel is the lateral view and the bottom panel is the dorsal view.

III. RESULTS

A. Echolocation signals

Forty-three files of echolocation clicks were collected during 4 days of measurements. Only echolocation events in which the amplitude of the signals received by the center hydrophone was within 3 dB of the highest were accepted for analysis. These criteria were used to ensure that a killer whale echolocation beam was directed at the array. The beam pattern measured for three different odontocete species (Au, 1993; Au *et al.*, 1995) indicate that when the major axis of the beam is directed to within $\pm 5°$ of the center hydrophone, the signal received by the center hydrophone will be within 3 dB of the highest amplitude. This 3-dB criterion is valid for ranges of approximately 5 to about 15-m range if a whale's head orientation was along the perpendicular bisect of the array. However, as the whales approached within about 5 to 10 m of the boat, they would totally submerge and swim to the side or below the array to avoid hitting it but still direct their signals towards the array. The whales almost always pointed their head towards the array as they approach the boat so that most of the echolocation clicks collected at short ranges were directed towards the center of the array.

Four spectra of echolocation click trains are shown in Fig. 4 presented in a waterfall format. The spectra suggest that portions of the signals in a click train can be relatively stable in shape, but also include portions that are highly variable and complex. Most of the energy in the spectra is between 20 and 60 kHz, much lower than for spectra of echolocation signals measured for smaller odontocetes in the field (Au and Herzing, 2003; Rasmussen *et al.*, 2002). Most of the clicks had bimodal frequency spectra with both primary and secondary peaks separated by as little as 10–15 kHz. The secondary peaks vary in shape from a slight ''bump'' in spectra to a clearly defined peak. The frequency of the bumps and distinct peaks were usually at similar frequencies.

Three representative types of clicks are shown in Fig. 5 with the signal waveforms on the right and the frequency spectra on the left. These clicks are very brief, generally between 80 and 120 μs in duration, with broad frequency spectra. The presence of clicks with bimodal spectra is obvious in the spectra plots. The majority (~89%) of the clicks examined had bimodal spectra. The killer whale click waveforms resemble those of other odontocetes such as *Tursiops truncatus, Delphinapterus leucas* (Au, 1993), *Pseudorca crassidens* (Au *et al.*, 1995), *Lagenorhynchus albirostris* (Rasmussen *et al.*, 2002), and *Grampus griseus* (Phillips *et al.*, 2003), but had almost twice as long duration and were almost an octave lower in frequency. Signals with a single high-frequency peak usually associated with higher amplitude signals (Au *et al.*, 1995) were rarely observed.

Orcinus projects relatively high amplitude signals with the maximum peak-to-peak source level measured at about 224 dB re:1 μPa. Most of the signals (75%) had source levels between 195 and 210 dB re:1 μPa. The peak-to-peak source level as a function of range between an echolocating

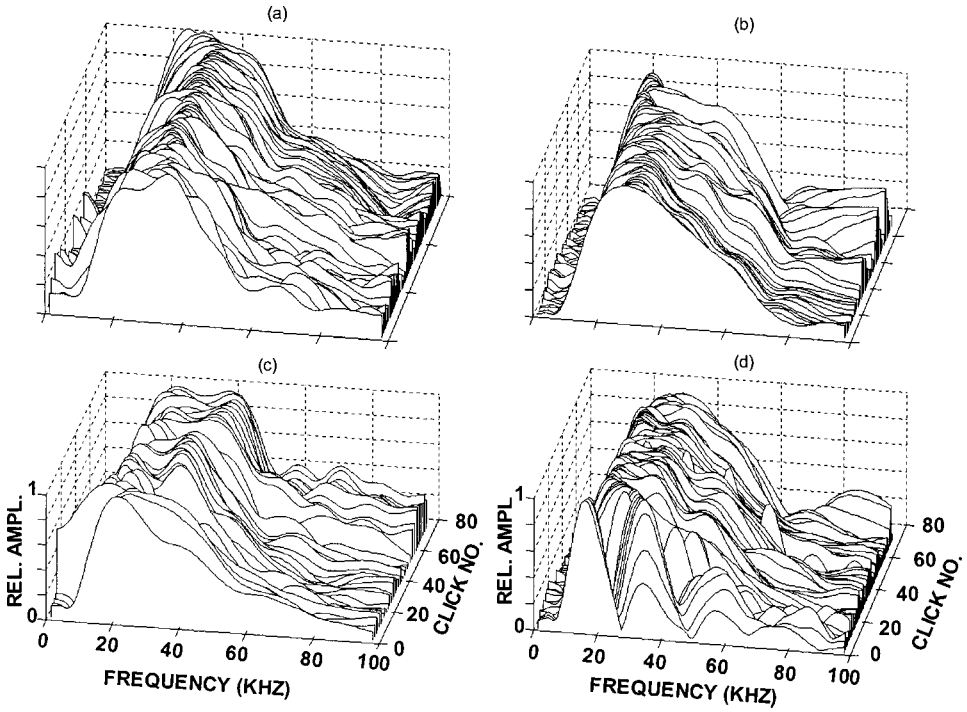


FIG. 4. Examples of the frequency spectra of four *Orcinus orca* echolocation click trains in a waterfall format.

killer whale and the array is shown in Fig. 6. As the whale's range to the array decreased, the source level also decreased. The solid curve in Fig. 6 is a regression curve represented by the equation

$$\mathrm{SL(dB)} = 181.4 + 20\log(R), \tag{2}$$

and has an r^2 value of 0.56, where R is the range in meters. The decrease in source level corresponded to the decrease in the one-way spherical spreading loss. Therefore, the amplitude of the echoes returning to the whales increased in magnitude as the range decreased, suggesting that killer whales prefer to receive echoes that have increasing signal to noise. Figure 6 also suggests that the whales were echolocating on the hydrophone array and not on some other objects, since the source level decreased as the range to the array decreased.

The histogram of center frequency is shown in Fig. 7. Center frequency is defined as that frequency which divides the energy in a frequency spectrum into two equal parts. The peak in the distribution of center frequency shown in the histogram occurred at 50 kHz, with 76% of the signals having center frequencies greater than 50 kHz. For bimodal spectra, center frequency is a better description of the fre-

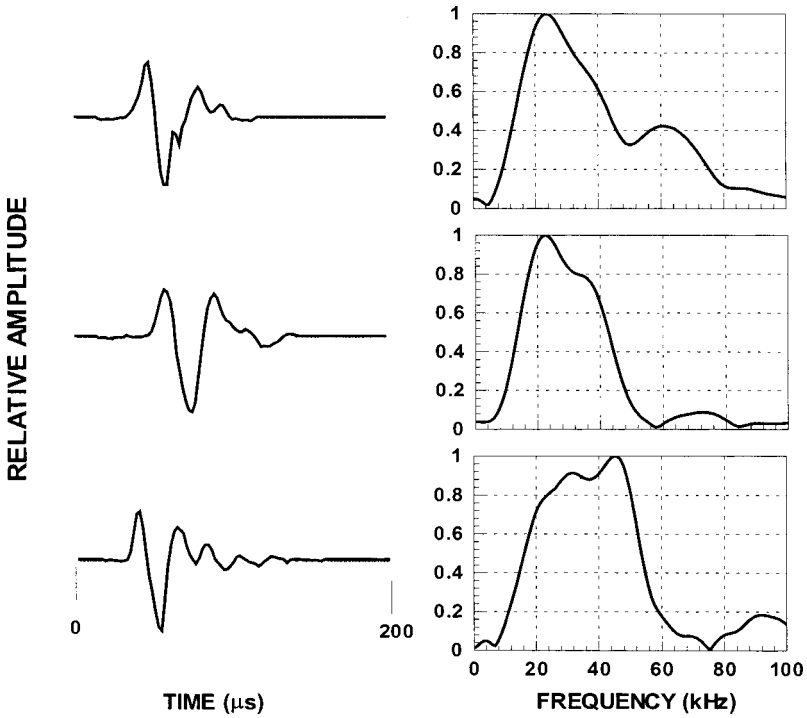


FIG. 5. Examples of the some representative waveforms and frequency spectra emitted by *Orcinus orca* in Johnstone Strait, Canada.

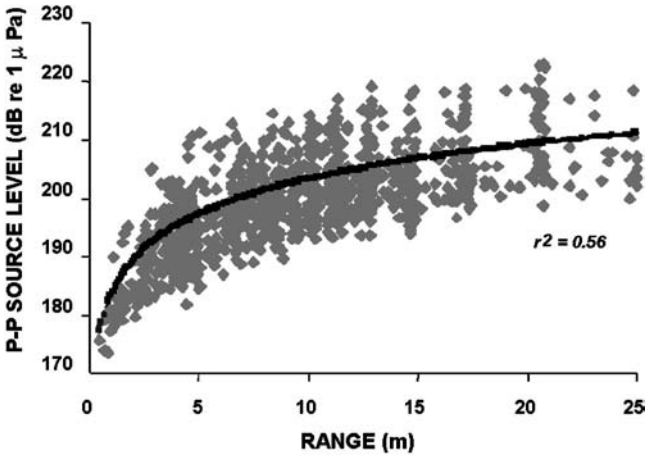


FIG. 6. Peak-to-peak source level as a function of range. The source levels are depicted by the gray diamond symbols. The solid line is the best-fit regression curve represented by the equation $A + 20 \log R$, where A is a constant and R is the range between a killer whale and the symmetrical star array.

quency spectra than peak frequency (Au *et al.*, 1995). Eighty-nine percent of the echolocation clicks had bimodal spectra in which the amplitude of the secondary peaks were within 50% of the amplitude of the primary peak.

The histogram of rms bandwidth is shown in Fig. 8. The rms bandwidth is a measure of the frequency width of a spectrum about the center frequency. The 3-dB bandwidth for bimodal spectra can often provide a misrepresentation of the width of the signal since the bandwidth might cover only the frequency range about the peak frequency. The rms bandwidth is probably a better measure of the width of signals with bimodal spectra. Most of the signals (83%) had an rms bandwidth between 35 and 50 kHz, with the peak in the histogram at 40 kHz. The Q of a signal is defined as the ratio of the center frequency over the rms bandwidth, and is equal to 0.9 to 1.4 for our data. These Q values indicate that *Orcinus* projects broadband echolocation signals.

B. Salmon target strength

The salmon was modeled using 775 fluid- and 290 gas-filled cylinders representing the body and swimbladder, respectively. Predicted target strength values plotted in three

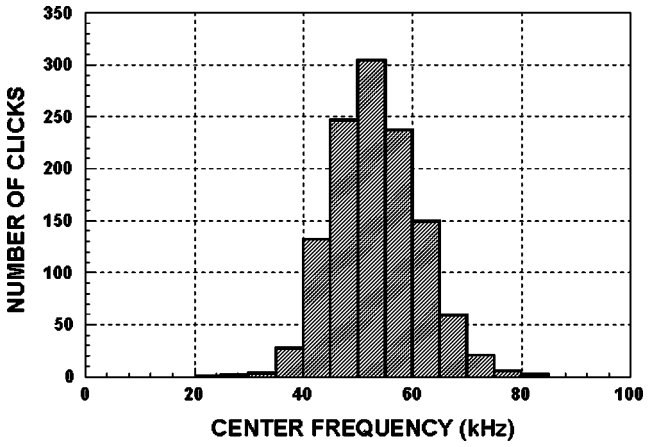


FIG. 7. Histogram of center frequency of echolocation signals.

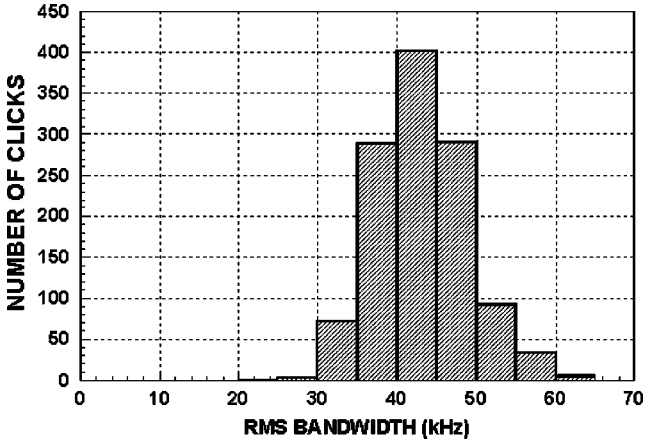


FIG. 8. Histogram of rms bandwidth of echolocation signals.

dimensions as a function of the tilt angle θ and the roll angle φ are shown in Fig. 9. KRM-predicted backscatter intensities in the killer whale ''attack quadrant'' (90° tilt, 180° roll) depend on the orientation of the fish relative to the incident echolocation acoustic wave. Target strengths vary by approximately 40 dB within the attack quadrant, increasing from low values at tail-on incidence to peak values at dorsal incidence. In the roll plane, maximum backscatter occurs at lateral incidence. The ridges and folds are due to constructive and destructive interference between the body and swimbladder at 50 kHz.

The target strength along the plane that is passed through the 0° roll of Fig. 9 is shown in Fig. 10. The degree values in Fig. 10 refer to the incident angle θ in the tilt plane, which matches the angle shown in Fig. 1. Also shown in the figure is the five-point moving average of target strength at 50 kHz over a range of 30° to 85°. The five-point moving average process effectively smoothened the target strength

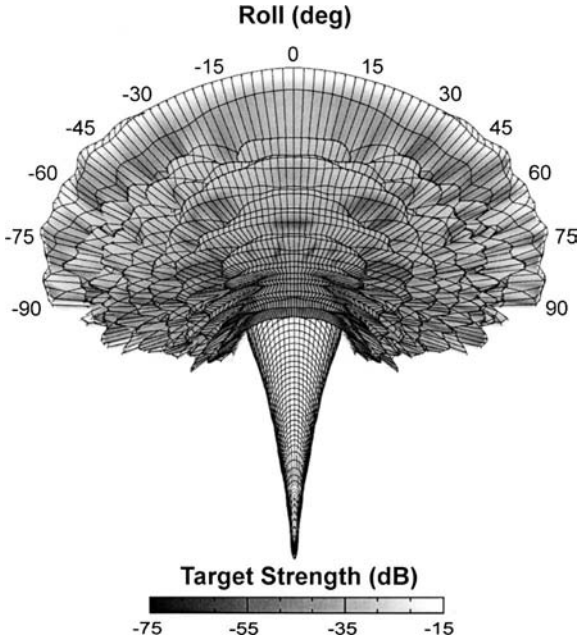


FIG. 9. Predicted Chinook salmon (*Oncorhynchus tshawytscha*) target strength (dB) at 50 kHz viewed at a 45° angle looking toward the head of the fish. KRM backscatter predictions are resolved at 2° in the tilt plane and 2° in the roll plane. Backscatter amplitude is depicted as distance from the center of the fish and using a gray scale where dark is low and white is high.

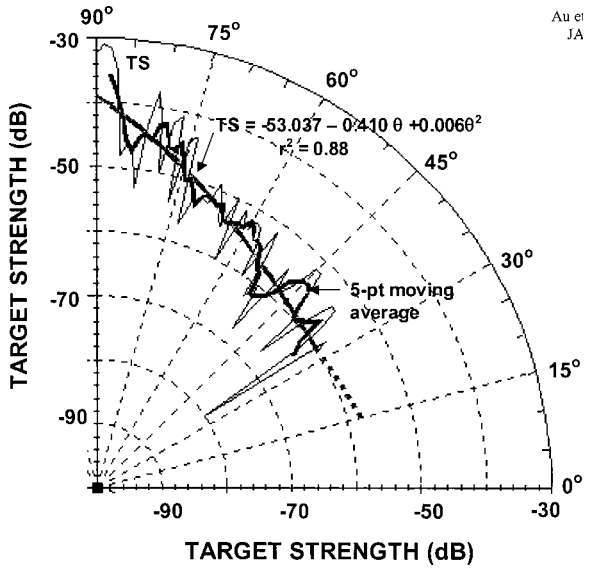


FIG. 10. Polar plot of chinnook salmon (*Oncorhynchus tshawytscha*) predicted target strength in the tilt plane along its spine at a 0° roll angle. The thin, black line is the KRM predicted backscatter at 50 kHz. The thick, black line is a five-point moving average of the target strength, starting at 30°. The gray solid curve is a second-order polynomial curve fit to the moving average.

results and removed the deep nulls. Target strength measurements using broadband, click signals indicate that deep nulls such as those observed in Fig. 10 do not exist (Benoit-Bird *et al.*, 2003). As a killer whale pursues and closes in on a prey, the attack angle between the predator and a prey will continuously change along with the target strength of the prey. A second-order polynomial was fitted to the 5-point moving average target strength curve ($r^2=0.88$). The equation for the fitted curve is

$$\mathrm{TS}(\theta) = -53.037 - 0.410 + 0.006\,\theta^2, \tag{3}$$

where θ is the attack angle in degrees and is defined by Eq. (1) depicted in Fig. 1.

C. Calculation of echo level

The level of the echoes (EL) in dB reflecting off a salmon and arriving back to the killer whale can be calculated with the equation

$$\mathrm{EL} = \mathrm{SL} - \mathrm{TL} + \mathrm{TS}, \tag{4}$$

where SL is the peak-to-peak source level, TL is the two-way transmission loss, and TS is the target strength. Substituting Eq. (2) for the SL and Eq. (3) for TS, and assuming spherical spreading loss, the peak-to-peak echo level can be expressed as

$$\mathrm{EL} = 128.633 - 20\log R - 2\alpha R - 0.410\,\theta + 0.006\,\theta^2, \tag{5}$$

where α is the absorption loss at 50 kHz and is approximately 0.016 dB/m (Urick, 1983). The echo level as a function of horizontal range between the salmon and the killer whale is plotted in Fig. 11 for a salmon depth of 25, 45, and 65 m. The echo levels for different salmon depths are not very different; at 100 m there is only a 4-dB difference between the salmon depth of 25 and 65 m. The differences become smaller for shorter horizontal ranges.

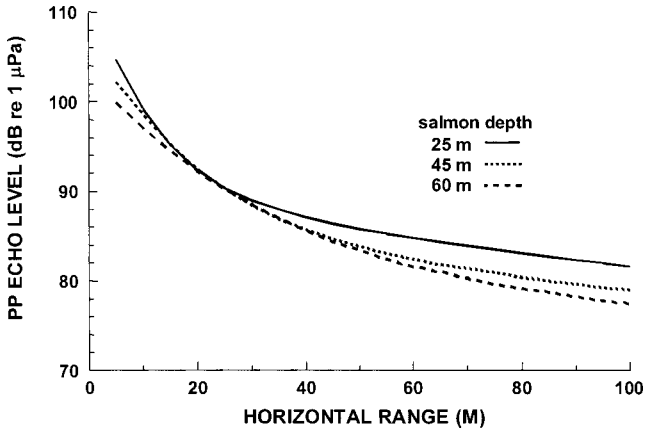


FIG. 11. Estimated peak-to-peak echo level of backscatter from a 0.7-m chinook salmon swiming at a depth of 30–50 m as a function of the horizontal range between the salmon and a foraging killer whale.

D. Detectability in quiet and noisy conditions

The echo levels shown on Fig. 11 can be compared with hearing threshold data for *Orcinus* obtained with a relatively broadband acoustic stimuli. This is important because hearing threshold obtained with narrow-band tonal signals may be different than thresholds obtained with broader band signals (Au *et al.*, 2002). Szymanski (1999) determined the hearing threshold of a killer whale by measuring the auditory brainstem response (ABR) using short tone bursts having durations of 0.5 ms as the acoustic stimuli. Such a short tone burst will have a broadband characteristic. ABR thresholds of 48 dB *re* 1 μPa peak-to-peak, for a 45-kHz tone burst and 52 dB for a 60-kHz tone burst were measured. Using a linear interpolation, the ABR threshold at 50 kHz should be approximately 49 dB. At a horizontal range of 100 m between a killer whale and a chinook salmon prey, the echo levels should vary between 82 and 78 dB for a salmon depth of 25 and 65 m, respectively. Therefore, in a quiet environment an echolocating killer whale would receive echoes that are between 29 and 33 dB above the ABR threshold at a horizontal range of 100 m. The echo levels shown in Fig. 11 taper off slowly, with range beyond 100 m suggesting the detection range probably extends considerably beyond 100 m. At the present time it is uncertain what would be considered the maximum range at which a killer whale would attack a prey, so we hesitate to extrapolate to longer ranges.

Under conditions in which an *Orcinus*' hearing is masked by ambient noise, the received echo level will be the same as in Fig. 11, but the received noise must also be considered in order to estimate detection ranges. The amount of noise (NL) that an animal will receive into its auditory system can be expressed as

$$\mathrm{NL} = N_0 + \mathrm{BW} - \mathrm{DI}, \tag{6}$$

where N_0 is the noise spectral density of the ambient noise, BW is the received bandwidth, and DI is the directivity index of the animal's auditory system. The receive bandwidth is a difficult parameter to deal with since the transmitted signal is

broadband, so that echoes produced by the signal will also be broadband and probably cover several critical bands of the animal's auditory system. One way of overcoming this difficulty is to use the rms bandwidth of the transmitter signal. From the histogram of Fig. 8, most of the signals had an rms bandwidth between 35 and 50 kHz. If we choose a 40-kHz bandwidth, then the bandwidth in dB will be

$$\mathrm{BW} = 10\log(40\ \ \mathrm{kHz}) = 46\ \ \mathrm{dB}. \tag{7}$$

The only odontocete in which the receiving beam pattern has been measured and the receiving directivity index estimated is *Tursiops* (Au and Moore, 1984). At 50 kHz, DI is approximately equal to 11 dB. The receiving directivity index for a killer whale can be estimated by modeling the sound-receiving system of both species using a circular disk transducer. The receiving directivity index for a circular transducer at a specific frequency is proportional to the radius, i.e., the larger the radius the larger the directivity index. Using the same line of reasoning with the killer whale, its receiving directivity index will be larger than that of the bottlenose dolphin (Au *et al.*, 1999). Therefore, the directivity index for a killer whale can be estimated by using the equation

$$\mathrm{DI_{OO}} = \mathrm{DI_{TT}} + 20\log\left(\frac{a_{\mathrm{OO}}}{a_{\mathrm{TT}}}\right), \tag{8}$$

where $\mathrm{DI_{OO}}$ is the directivity index for *Orcinus*, $\mathrm{DI_{TT}}$ is the directivity index for *Tursiops*, a_{OO} is the radius of an killer whale's head, and a_{TT} is the radius of a *Tursiops* head measured at the blow hole. The ratio of $a_{\mathrm{OO}}/a_{\mathrm{TT}}$ varies between 3 and 3.5. If we use a ratio of 3, the directivity index for an *Orcinus* at 50 kHz will be approximately 10 dB larger than that of *Tursiops*, or 21 dB. Let us assume a sea state 4 condition, in which the noise spectral density at 50 kHz is approximately $35\,\mathrm{dB}\,re{:}\,1\ \mu\mathrm{Pa}^2/\mathrm{Hz}$ (Urick, 1983). Therefore, from Eq. (6) the received noise level centered at 50 kHz will be approximately 60-dB rms. Since the source level in Eq. (4) and the ABR threshold are given in terms of peak-to-peak sound-pressure levels, the received noise should also be expressed in terms of its peak-to-peak value. Using the relationship between the rms and peak-to-peak values of a sine wave, we merely add 9 dB to the rms value and obtain a peak-to-peak noise level in the frequency band centered at 50 kHz of 69 dB. Therefore, even in sea state 4, the echo level from a chinook salmon would be at least 9 dB above the received noise level at a horizontal range of 100 m.

Another source of ambient noise besides wind and wave is falling rain. Rain can increase the ambient noise level considerably, depending on the rate of precipitation and sea state conditions (Scrimger *et al.*, 1989). For rain fall of 3 mm/hr, the rms noise spectral level at 50 kHz is approximately 42 dB, which is only 7 dB above sea state 4 noise level (Urick, 1983). The received echo should be at least 2 dB above the rain noise at a horizontal range of 100 m and increase rapidly for shorter horizontal ranges. However, for moderately heavy rain the rms noise spectral level will increase to approximately $52\,\mathrm{dB}\,re{:}\,1\ \mu\mathrm{Pa}^2/\mathrm{Hz}$ (Urick, 1983) so that the horizontal range must be shorter than about 40 m before the signal begins to extend above the noise level. The situation will not be as drastic as it seems since rain noise originates at the surface, and if the *Orcinus*' beam is pointed below the horizon the received noise will be less than predicted by Eq. (6). The use of the directivity index in Eq. (6) assumes that the noise is isotropic (Urick, 1983) and can only be used to obtain a rough estimate of the received anisotropic rain noise. Nevertheless, under moderately heavy rain conditions, the detection range of *Orcinus* may be reduced significantly but probably not enough to prevent it from foraging by echolocation. Furthermore, the whale might change hunting tactics and swim some depth below the surface in order to avoid the rain noise at the surface.

IV. DISCUSSION AND CONCLUSIONS

Orcinus projects broadband, short-duration echolocation signals similar to those of other odontocetes. The signals are brief, with only three to seven oscillations, and broadband with a Q between 0.9 and 1.4. The signals have a duration that is approximately twice as long as those used by *Tursiops* and other small odontocetes (Au, 1993), with a center frequency lower by close to one octave. Most of the signals have a bimodal frequency distribution which contributes to the broadband nature of the signals. The bimodal distribution has a low-frequency peak between 20 and 30 kHz and a high-frequency peak between 40 and 60 kHz, with the peak in the center frequency histogram at 50 kHz. The broad bandwidth of the echolocation signal provides a good range resolution capability (Au, 1993) that should enable *Orcinus* to perform fine target discrimination, especially close to where the rocky cliff protrudes steeply into the water in Johnstone Strait. Salmon have been known to seek refuge along the cliffs in this area from foraging killer whales (Ford *et al.*, 1998).

The peak-to-peak source levels measured for *Orcinus* are comparable to those measured for *Tursiops* in open-water captive echolocation experiments (Au, 1993; Au and Snyder, 1980) for comparable target ranges. For target ranges between 6 and 20 m, *Tursiops* source levels varied from about 204 to $216\,\mathrm{dB}\,re{:}\,1\ \mu\mathrm{Pa}$, which is similar to that of *Orcinus*. However, there is a large difference between the target strength of targets used in the echolocation experiments with *Tursiops* and the target strength of the array assembly used to measure signals in the field. Although the target strength of the array was not measured, the theoretical target strength of an aluminum pipe, connected to a flat Plexiglas container mounted on a flat delrin plate along with a camera holder, should be approximately 15–20 dB greater than the small cylinders and spheres used in the *Tursiops* experiments (Au, 1993; Au and Snyder, 1980). The fact that the amplitude of echolocation signals used by free-ranging dolphins is comparable to those used by captive dolphins suggests the importance of range on the source levels utilized by dolphins. Despite the higher target strength of the array, the killer whale emitted similar source levels as *Tursiops* echolocating on much weaker targets at similar ranges.

The amplitude of the echolocation signals is sufficiently high for the killer whales to detect chinook salmon prey at relatively long ranges ($\approx$100 m) in the presence of relatively high wind and wave conditions. Refined estimates of echo

levels and detection ranges can be obtained by further research on the echolocation and auditory capabilities of *Orcinus*, and on broadband backscatter from salmon and other fishes. Some parameters that need to be obtained for *Orcinus* include the echolocation detection threshold in quiet and in noise, receiving directivity index, and receiving bandwidth applicable to echolocation. Models of acoustic backscatter from fishes with a swimbladder for broadband, short-duration signals also need to be developed. The echo levels obtained in this study are based on the estimates of missing or unknown parameters pertaining to both the killer whale and salmon. Despite the presence of certain unknowns, it is doubtful that the results of this study would be significantly altered if all the parameters were known. The echo level at a horizontal range of 100 m is sufficiently high that slight adjustments in the echo level would probably not have any consequence. The pursuit geometry in which a prey is swimming directly away from the predator used in this study resulted in relatively low target strength values for the chinook salmon. Higher target strength can be expected for a geometry in which the longitudinal axis of a salmon is close to being perpendicular to the trajectory of the echolocation signal emitted by a killer whale. Here, we assume that the *Orcinus*' head is oriented so that the acoustic axis of its sonar system is pointed directly at the prey.

An interesting issue that has arisen as the result of this study is associated with the range at which a killer whale predator would consider a chinook salmon as a prey worth pursuing. The amount of energy that a predator must expend in order to close the distance between it and the prey can be assumed to be directly proportional to the distance between it and the prey. Just because an *Orcinus* can detect a chinook salmon at ranges in excess of 100 m does not necessarily mean that it will start pursuing that salmon from that range. There is a good possibility that the predator would continue to hunt for a prey that would be at a closer range. It is also likely that the whale may continue swimming in the direction of a salmon until the separation distance closes to a particular limit before it begins a pursuit of the prey. Unfortunately, little is known on this important and interesting topic.

There is also the issue of the prey possibly hearing the echolocation signals of a predator killer whale. Unfortunately, there is little information on the hearing characteristics of salmon. Hawkins (1978) measured the audiogram of an Atlantic salmon (*Salmo salar*) and found an upper frequency of hearing of approximately 400 Hz. We should not expect much difference in the upper limit of hearing for other salmon species. Therefore, we could surmise that it would be very unlikely for a Chinook salmon to detect the echolocation signals of a foraging killer whale.

The foraging behavior of *Orcinus* in the vicinity of Vancouver Island pursuing individual salmon is in contrast to the foraging behavior of killer whales feeding on herring in Norway (Nottestad *et al.*, 2002). A herring school would present a much larger target strength than an individual salmon. The target strength of a school would probably not depend on the direction from which a sonar signal arrives. Therefore, the detection range should be greater with a school and it would not depend strongly on the orientation of the killer whale with respect to the school.

This is the first study to model prey detection by echolocation for any odontocete. Although certain assumptions associated with the killer whale's echolocation and auditory systems were made, based on studies on the Atlantic bottlenose dolphins, we feel that the results obtained are representative of the *Orcinus orca*. Our results suggest that killer whales foraging on chinook salmon in coastal waters of British Columbia are well equipped to detect their prey at ample ranges to be successful at the task.

ACKNOWLEDGMENTS

We thank James and Mary Borrowman of Stubbs Island Whale Watching, Telegraph Cove, British Columbia, for use of their skiff and for hosting us. We also thank Graeme Ellis for the use of his boat and for his assistance. This study was supported in part by the Office of Naval Research, Grant No. N00014-99-1-0800 to Paul E. Nachtigall and Whitlow Au, from Dr. Robert Gisiner, Program Manager. Support for the fish backscatter research came from ONR's Biological/Chemical Oceanography, Ocean Acoustics Science Divisions. The support of the Vancouver Aquarium Marine Science Centre is also appreciated. This is HIMB contribution No. 1170.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer, New York).

Au, W. W. L., and Herzing, D. L. (**2003**). ''Echolocation signals of wild Atlantic spotted dolphin (*Stenella frontalis*),'' J. Acoust. Soc. Am. **113**, 598–604.

Au, W. W. L., Kastelein, R. A., Ripper, T., and Schooneman, N. M. (**1999**). ''Transmission beam pattern and echolocation signals of a harbor porpoise (*Phocoena phocoena*),'' J. Acoust. Soc. Am. **106**, 3699–3705.

Au, W. W. L., Lemonds, D. W., Vlachos, S., Nachtigall, P. E., and Roitblat, H. L. (**2002**). ''Atlantic bottlenose dolphin hearing threshold for brief broadband signals,'' J. Comp. Psychol. **116**, 151–157.

Au, W. W. L., and Moore, P. W. B. (**1984**). ''Receiving beam patterns and directivity indices of the Atlantic bottlenose dolphin, *Tursiops truncatus*,'' J. Acoust. Soc. Am. **75**, 255–262.

Au, W. W. L., Pawloski, J. L., Nachtigall, P. E., Blonz, M., and Gisner, R. C. (**1995**). ''Echolocation signals and transmission beam pattern of a false killer whale (*Pseudorca crassidens*),'' J. Acoust. Soc. Am. **98**, 51–59.

Au, W. W. L., and Snyder, K. J. (**1980**). ''Long-range target detection in open waters by an echolocating Atlantic bottlenose dolphin,'' J. Acoust. Soc. Am. **56**, 1280–1290.

Aubauer, R. (**1995**). *Korrelationsverfahren zur Flugbahnverfolgung echoortender Fledemause*, Unpublished Dipl.-Ing, Technishen Hochschule Darmstadt, Darmstadt.

Barret-Lennard, L. G., Ford, J. K. B., and Heise, K. A. (**1996**). ''The mixed blessing of echolocation: Differences in sonar use by fish-eating and mammal-eating killer whales,'' Anim. Behav. **51**, 553–565.

Benoit-Bird, K. J., Au, W. W. L., and Kelley, C. D. (**2003**). ''Acoustic backscattering by Hawaiian Lutjanid snappers. I. Target strength and swimbladder characteristics,'' J. Acoust. Soc. Am. **114**, 2757–2766.

Candy, J. R., and Quinn, T. P. (**1999**). ''Behavior of adult chinook salmon (*Oncorhynchus tshawytscha*) in British Columbia coastal waters determined from ultrasonic telemetry,'' Can. J. Zool. **77**, 1161–1169.

Clay, C. S., and Horne, J. K. (**1994**). ''Acoustic models of fish: The Atlantic cod (*Gadus morhua*),'' J. Acoust. Soc. Am. **96**, 1661–1668.

Ford, J. K. B., Ellis, G. M., and Balcomb, K. C. (**2000**). *Killer Whales: The Natural History and Genealogy of Orcinus orca in British Columbia and Washington* (UBC Press and U. of Washington Press, Vancouver, B.C., and Seattle, WA).

Ford, J. K. B., Graeme, E. M., Barrett-Lennard, L. G., Morton, A. B., Palm, R. S., and Balcomb III, K. C. (**1998**). ''Dietary specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British Columbia and adjacent waters,'' Can. J. Zool. **76**, 1456–1471.

Hawkins, A. D., and Johnstone, A. D. F. (**1978**). ''The hearing of the Atlantic salmon *Salmo salar*,'' J. Fish Biol. **13**, 655–673.
Horne, J. K., and Jech, J. M. (**1999**). ''Multi-frequency estimates of fish abundance: Constraints of rather high frequencies,'' ICES J. Mar. Sci. **56**, 1054–1199.
Jech, J. M., and Horne, J. K. (**2002**). ''Three-dimensional visualization of fish morphometry and acoustic backscatter,'' ARLO **3**, 35–40.
Jech, J. M., Schael, D. M., and Clay, C. S. (**1995**). ''Application of three sound scattering models to threadfin shad (*Dorosoma petenense*),'' J. Acoust. Soc. Am. **98**, 2262–2269.
Kellogg, W. N. (**1958**). ''Echo ranging in the porpoise,'' Science **128**, 982–988.
Nachtigall, P. E., and Moore, P. W. B. (Eds.) (**1988**). *Book Animal Sonar: Processes and Performance* (Plenum, New York).
Nottestad, L., Misund, O. A., Ferno, A., Melle, W., Mackinson, S., Axelsen, B., and Pitcher, T. (**2002**). ''Synoptic field studies of Norwegian spring-spawning herring: Acoustics, Behaviour and Ecology,'' Paper presented at the ICES Symposium of Acoustics in Fisheries and Aquatic Ecology, Montpellier, France, 10–14 June, 2002.
Phillips, J. D., Nachtigall, P. E., Au, W. W. L., Pawloski, J. L., and Roitblat, H. L. (**2003**). ''Echolocation in Risso's dolphin (*Grampius griseus*),'' J. Acoust. Soc. Am. **113**, 605–616.
Rasmussen, M. H., Miller, L. A., and Au, W. W. L. (**2002**). ''Source levels of clicks from free-ranging white beaked dolphins (*Lagenorhynchus albirostris* Gray 1846) recorded in Icelandic waters,'' J. Acoust. Soc. Am. **111**, 1122–1125.
Scheville, WE, and Lawrence, B. (**1953a**). ''Auditory response of a bottlenose porpoise, Tursiops turncatus, to frequencies above 100 Kc.'' J. Exp. Zool. **124**, 147–165.
Scheville, W., and Lawrence, B. (**1953b**). ''High-frequency auditory response of a bottlenosed porpoise, *Tursiops truncatus*, (Montague).'' J. Acoust. Soc. Am. **25**, 1016–1017.
Scrimger, J. A., Evans, D. J., and Yee, W. (**1989**). ''Underwater noise due to rain—open ocean measurements,'' J. Acoust. Soc. Am. **85**, 726–731.
Szymanski, M. D., Bain, D. E., Kiehi, K., Pennington, S., Wong, S., and Henry, K. R. (**1999**). ''Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioral audiograms,'' J. Acoust. Soc. Am. **106**, 1134–1141.
Urick, R. J. (**1983**). *Principles of Underwater Sound* (McGraw-Hill, New York).

Old World frog and bird vocalizations contain prominent ultrasonic harmonics

Peter M. Narins[a)]
Departments of Physiological Science and OBEE Biology, University of California, 621 Charles E. Young Drive S., Los Angeles, California 90095

Albert S. Feng and Wenyu Lin
Department of Molecular and Integrative Physiology, University of Illinois, Urbana, Illinois 61801

Hans-Ulrich Schnitzler and Annette Denzinger
Lehrstuhl Tierphysiologie, D-72076 Tübingen, Germany

Roderick A. Suthers
Medical Sciences and Department of Biology, Indiana University, Bloomington, Indiana 47405

Chunhe Xu
Shanghai Institutes of Biological Sciences, The Chinese Academy of Science, Shanghai, P. R. China

(Received 14 October 2003; revised 5 November 2003; accepted 6 November 2003)

Several groups of mammals such as bats, dolphins and whales are known to produce ultrasonic signals which are used for navigation and hunting by means of echolocation, as well as for communication. In contrast, frogs and birds produce sounds during night- and day-time hours that are audible to humans; their sounds are so pervasive that together with those of insects, they are considered the primary sounds of nature. Here we show that an Old World frog (*Amolops tormotus*) and an oscine songbird (*Abroscopus albogularis*) living near noisy streams reliably produce acoustic signals that contain prominent ultrasonic harmonics. Our findings provide the first evidence that anurans and passerines are capable of generating tonal ultrasonic call components and should stimulate the quest for additional ultrasonic species. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1636851]

PACS numbers: 43.80.Ka, 43.80.Gx, 43.80.Nd, 43.66.Gf [WA] Pages: 910–913

I. INTRODUCTION

Animal vocal signals are extraordinarily diverse. Their sound intensities range widely, and their dominant frequencies may extend from below the human hearing range (infrasounds) to a few Hertz through the audible range (20 Hz to 20 kHz) and into the ultrasonic range (to >200 kHz). It is thought that coevolution of the vocal and auditory systems results in the generation of sound communication signals that are correlated with the ability of an individual (or a conspecific individual) to detect and perceive that signal. Nevertheless, the hearing range of an animal need not coincide precisely with the frequency range of its vocalizations (Konishi, 1971). The logical outcome is a wide diversity of auditory processing abilities among animals (Stebbins, 1983). Common exceptions to this correlation are impulsive, click-like signals that are broadband, often containing frequencies exceeding the auditory system's upper range of detectable signals.

Most anuran amphibians (frogs and toads) produce advertisement calls containing frequency components in the range from ~ 100 Hz to 5–6 kHz (Capranica, 1965; Glaw and Vences, 1994). In contrast, birdsong may contain frequencies as low as 80–90 Hz and harmonic components as high as 10.7 kHz, a range of about seven octaves (Greenwalt, 1968). *Amolops tormotus* is an arboreal frog in the family Ranidae restricted in its distribution to two provinces in central China (Zhou and Adler, 1993). Males of this species call nightly from the low vegetation along the banks of rivers and streams. Their vocal repertoire is extraordinarily rich; individual calls exhibit multiple upward and downward frequency sweeps, rapid frequency ''steps'' and sudden onset and offset of selective harmonic components within a call note (Feng *et al.*, 2002). *Abroscopus albogularis* (family Sylviidae), a leaf warbler distributed from Nepal to S. China and N. Indochina, inhabits evergreen forests and bamboo thickets in the same regions at altitudes up to 2440 m (Meyer de Schauensee, 1984). Its call has been described as ''a shrill twitter'' (MacKinnon and Phillipps, 2000). Both the frog, *A. tormotus*, and the songbird, *A. albogularis*, call from the vegetation along creeks and streams in Anhui Province, China. During the rainy season the water level rises precipitously, resulting in increased background noise levels in the vicinity of the streams (Feng *et al.*, 2002). In the present study, the advertisement call of the frog *A. tormotus* and the song of the songbird *A. albogularis* in their natural habitat were recorded using an ultrasonic microphone and a digital recording system to examine the vocal features of their vocalizations and their relationship to the stream noise.

II. METHODS

All recordings were made between 1900 and 2400 h along the Tau Hua Creek (30° 06′N, 118° 10′E) in Huang-

a)Electronic mail: pnarins@ucla.edu

shan, China on 10–17 May 2002. Animal vocalizations were recorded using a custom-built, PC-based recording device (PCTape), and a custom-made ultrasonic microphone (Department of Animal Physiology, University of Tubingen) with a flat (± 3 dB) frequency response from 15 to 120 kHz with a roll-off of 10 and 6 dB/oct at <15 kHz and >120 kHz, respectively. Signals were digitized using a 16-bit A/D converter (Analog Devices AD7723) at a sampling rate of 192 or 256 kHz, with $8\times$ oversampling. Data were saved as wave files and analyzed (FFT, 1024), and displayed using SELENA, a custom-designed program (S. Andrzheevski, St. Petersburg). Field comments were recorded with a separate microphone, which was fitted with a switch to allow remote control of PCTape.

III. RESULTS

We recorded and analyzed spontaneous vocalizations (not evoked by acoustic playbacks) from 40 frogs (*A. tormotus*) and two songbirds (*A. albogularis*) along the stream of the Tau Hua Creek. Measurement of the ambient noise generated by the running water with a precision sound level meter (GenRad 1982) with various filter settings showed that the noise was broadband with a peak near 100 Hz decaying by 19 dB at 2 kHz and by 63 dB at 28 kHz.

The frog, *Amolops*, typically produces two-note calls composed of short tone pips [Fig. 1(a)], or one-note calls having short [Figs. 2(b)and (c)] or long duration [Figs. 1(b), 1(c), and 2(a)]. These calls show varying degrees of frequency modulation (FM); sudden onset and offset of harmonic and subharmonic components are often observed. Notably, these calls display a distinct first formant in the ''audible range'' and a prominent second formant at ~ 60 kHz; in some calls there is a spectral trough around 32–45 kHz [Figs. 1(c), 2(a), and (c)]. Sometimes, a third formant can be observed at 105 kHz [Fig. 1(c)]. It is clear these calls have significant energy in the ultrasonic range. A fortuitous simultaneous recording of a passing bat and a two-note call of the frog, *A. tormotus*, reveals that many frequency components of the frog's call are shared with and even exceed those in the bat's frequency-modulated echolocation pulses [Fig. 3(a)]. Similar to *A. tormotus*, the pulsed song of the songbird, *A. albogularis* [Fig. 3(b)], is characterized by distinct formants in the audible and ultrasonic ranges, with harmonics up to 54 kHz [Fig. 3(b)].

IV. DISCUSSION

The vocal repertoire of males of the arboreal frog, *Amolops tormotus*, is extraordinarily rich and individual calls share many features of birdsong and primate vocalizations, such as multiple upward and downward FM sweeps, rapid frequency transitions, and multiple harmonics (Feng *et al.*, 2002). The previous and the present studies further show that their calls often exhibit nonlinear features, e.g., sudden onset and offset of harmonic and subharmonic components within a call note—these are unique among known amphibian vocalizations. In mammals and birds, nonlinear acoustic properties are generated by the larynx (Herzel *et al.*, 1995; Fitch *et al.*, 2002) and the syrinx (Fee *et al.*, 1998), respectively. The origin of nonlinearities in frogs has yet to be determined.

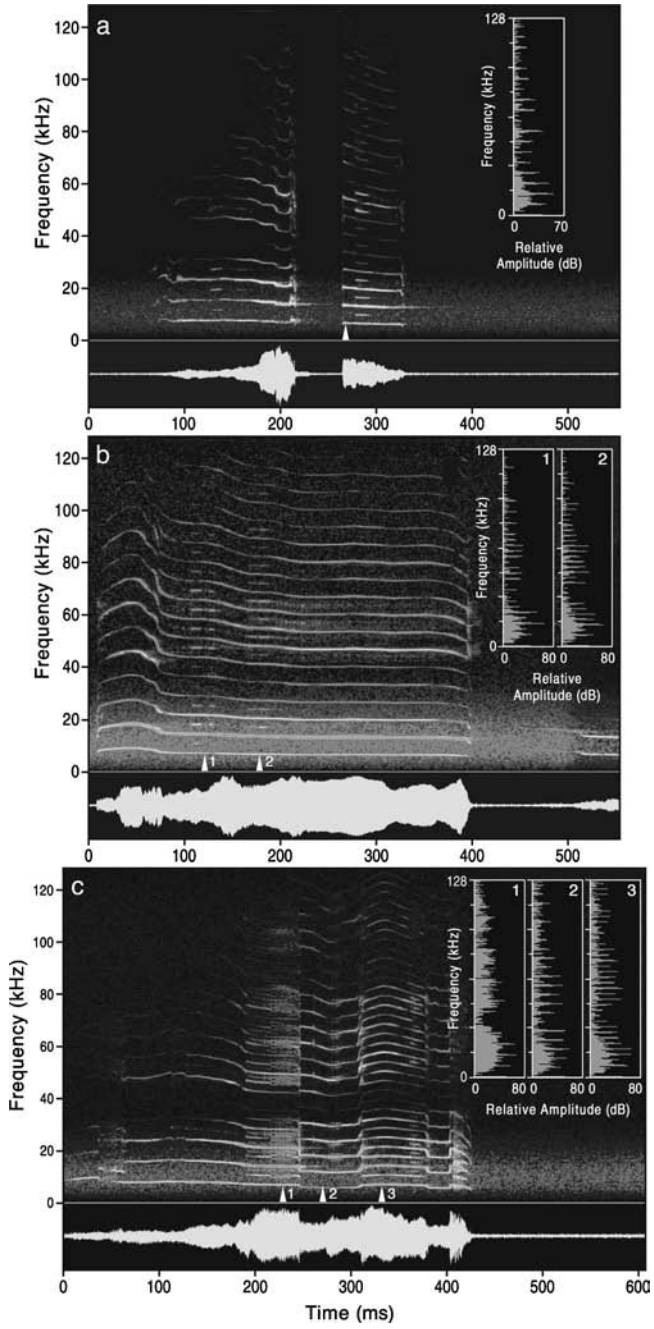


FIG. 1. Sound spectrograms (top panel), waveforms (bottom panel), and instantaneous amplitude spectra [insets, taken at indicated points (arrowheads)] of vocal signals of the frog, *Amolops tormotus* in Anhui Province, China. (a) A two-note call. The background noise (up to 20 kHz) is due to the rushing water in the Tau Hua Creek. Note that the roll-off of background noise energy below 15 kHz in the amplitude spectra is attributed to the limited-low frequency response of the ultrasonic microphone. (b)(c) One-note long calls of the frog, *A. tormotus*. The initial warble FM in (b) is followed by a constant-frequency tone with multiple harmonics; sudden onset of subharmonics can be seen in two instants (point 1, 2). For the one-note long call in (c), there is a sudden onset of subharmonics 200 ms into the call followed by a chaotic noisy segment (point 1) that itself ceases suddenly at 240 ms (point 2). For all plots, dynamic range: 90 dB; temperature range during recordings: 17–18 °C.

Published sound spectrograms of the calls of *A. tormotus* (Feng *et al.*, 2002) give a hint of the presence of ultrasonic

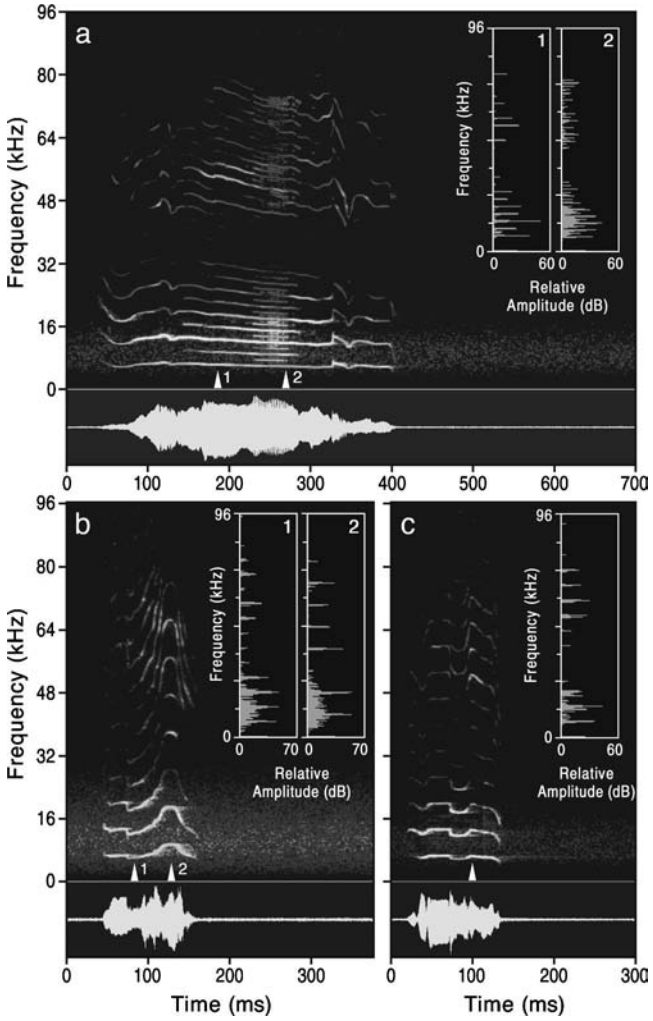


FIG. 2. Long call of the frog, *Amolops tormotus* (a), and two short calls [(b) and (c)] showing significant energy in the ultrasonic range and a spectral notch in the range of 32–45 kHz. Other data same as Fig. 1.

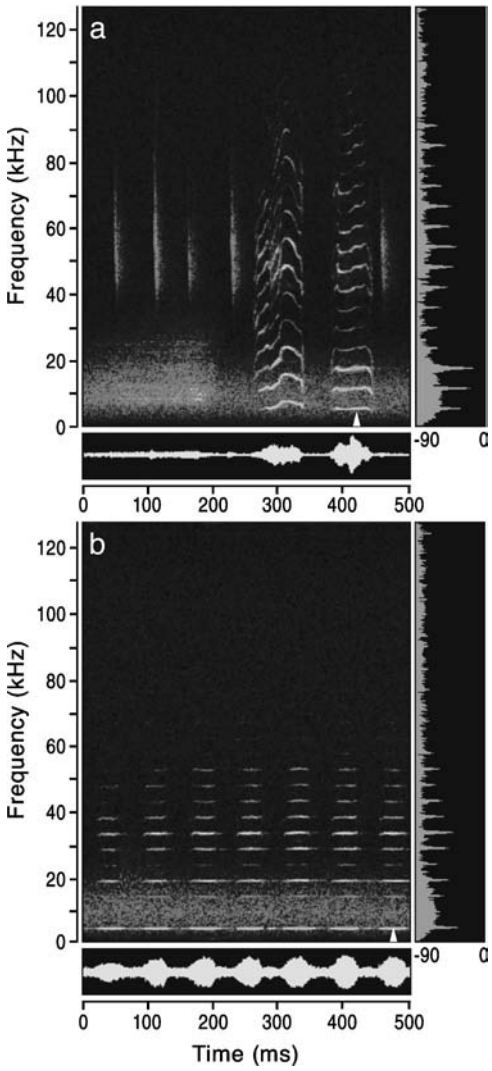


FIG. 3. Sound spectrograms (top left panel), waveforms (bottom panel), and average amplitude spectra [top right panel, 36-ms window centered at indicated points (arrowheads)] of animal vocalizations recorded at the study site. (a) Simultaneous recording of a short, two-note call of the frog, *A. tormotus*, and the FM echolocation pulses of a passing bat. (b) Pulsed song of the Rufous-faced warbler (*Abroscopus albogularis*); harmonics up to 54 kHz are evident. Other data same as Fig. 1.

components, but the recordings were low-pass filtered at 22 kHz by the frequency responses of the microphone (Sennheiser ME-66) and tape recorder (Sony TC-D5M) and by the limited frequency range of the sound analysis software (Canary version 1.2.4). In the present study, use of a wideband microphone and a high-speed digital recording system with an extended range up to 128 kHz allowed unambiguous documentation of the ultrasonic harmonic components in the vocalizations of both the frog, *A. tormotus*, and the songbird, *A. albogularis*.

Among the vertebrates, ultrasonic harmonics of acoustic signals are known to be produced only by mammals, including whales (Sales and Pye, 1974), dolphins (Au, 1993), bats (Griffin, 1958), and rodents (Ehret, 1992). Our findings provide the first evidence that anurans and passerines are capable of generating tonal ultrasonic call components and should stimulate the quest for additional ultrasonic species.

These results raise several questions: Why do these animals produce signals containing ultrasonic components? That is, what function do they perform? Well-trained opera singers are able to produce a distinct high-frequency formant near 3 kHz (known as the singer's formant) that allows their voices to stand out from the loud orchestra in the background (Sundberg, 1987). Thus, in this case, spectral salience may be viewed as an adaptation to overcome masking produced by the orchestra. In a recent study, Slabbekoorn and Peet (2003) showed that urban great tits (*Parus major*) in noisy locations sing with a higher minimum frequency than those in quieter locations, presumably avoiding masking of their songs by the predominantly low-frequency background noise. In the present study, we show that an Old World frog (*Amolops tormotus*) and an oscine songbird (*Abroscopus albogularis*) living near noisy streams produce acoustic signals that contain significant ultrasonic harmonics, which may also be the result of selective pressure to avoid masking by the wideband river noise that extends over the entire range of human hearing (Figs. 1 and 3). Alternatively, ultrasonic harmonics in the calls may be targeted at prey species or may be simply be an epiphenomenon—a by-product of the sound production mechanism. At present, our data do not allow us to distinguish between these alternatives.

Can *Amolops* or *Abroscopus* perceive these ultrasonic signal components? Current psychoacoustical data support an upper limit to frog and avian hearing of 5 kHz (Fay, 1988), and 12 kHz (Konishi, 1973), respectively. It must be noted, however, that systematic study of the response of the anuran or avian auditory system to ultrasound is lacking. In light of the results of the present study, it is important that the limits of hearing for these two groups of vertebrates be reevaluated.

ACKNOWLEDGMENTS

The authors wish to thank Chen Qi-Lin, Weng Jun, and Yu Xinjian for their invaluable assistance in the field, Nasser Farahbakhsh for his able assistance with analysis of the sound files, and Margaret Kowalczyk and Ingrid Kaipf for figure preparation. This work was supported by grants from the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health (R01DC00222 to PMN and R01DC00663 to ASF), and from a UCLA Academic Senate grant to PMN, a DFG grant (SFB 550) to HUS, and a grant from the State Key Basic Research and Development Plan [G1998010100 (China)] to CHX.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer-Verlag, New York).
Capranica, R. R. (**1965**). *The Evoked Vocal Response of the Bullfrog: A Study of Communication by Sound*. Res. Monograph 33 (MIT, Cambridge, MA).
Ehret, G. (**1992**). ''Categorical perception of mouse-pup ultrasounds in the temporal domain,'' Anim. Behav. **43**, 409–416.
Fay, R. R. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay, Winnetka, IL).
Fee, M. S., Shraiman, B., Pesaran, B., and Mitra, P. P. (**1998**). ''The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird,'' Nature (London) **395**, 67–71.
Feng, A. S., Narins, P. M., and Xu, C.-H. (**2002**). ''Vocal acrobatics in a Chinese frog, *Amolops tormotus*,'' Naturwissenschaften **89**, 352–356.
Fitch, W. T., Neubauer, J., and Herzel, H. (**2002**). ''Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production,'' Anim. Behav. **63**, 407–418.
Glaw, F., and Vences, M. (**1994**). *A Fieldguide to the Amphibians and Reptiles of Madagascar, 2nd ed.* (M. Vences & F. Glaw Verlags, Köln, Germany).
Greenwalt, C. H. (**1968**). *Bird Song: Acoustics and Physiology* (Smithsonian Institution, Washington, DC).
Griffin, D. (**1958**). *Listening in the Dark-The Acoustic Orientation of Bats and Men*. (Yale U.P., New Haven, CT).
Herzel, H., Berry, D., Titze, I., and Steinecke, I. (**1995**). ''Nonlinear dynamics of the voice: signal analysis and biomechanical modeling,'' Chaos **5**, 30–34.
Konishi, M. (**1971**). ''Comparative neurophysiological studies of hearing and vocalizations in songbirds,'' Z. Vergl. Physiol. **66**, 257–272.
Konishi, M. (**1973**). ''How the owl tracks his prey,'' Am. Sci. **61**, 414–424.
MacKinnon, J., and Phillipps, K. (**2000**). *A Field Guide to the Birds of China* (Oxford U.P., Oxford).
Meyer de Schauensee, R. (**1984**). *The Birds of China* (Smithsonian Institution, Washington, DC).
Sales, G. D., and Pye, J. D. (**1974**). *Ultrasonic Communication by Animals* (Chapman and Hall, London, UK).
Slabbekoorn, H., and Peet, M. (**2003**). ''Birds sing at a higher pitch in urban noise,'' Nature (London) **424**, 267.
Stebbins, W. C. (**1983**). *The Acoustic Sense of Animals* (Harvard U.P., Cambridge, MA).
Sundberg, J. (**1987**). *The Science of the Singing Voice* (Northern Illinolis U.P., Dekalb, IL).
Zhou, E. M., and Adler, K. (**1993**). *Herpetology of China* (Society for the Study of Amphibians and Reptiles, Oxford, Oxford, UK).

Click sounds produced by cod (*Gadus morhua*)

Heike I. Vester, Lars P. Folkow, and A. S. Blix
Department of Arctic Biology and Institute of Medical Biology, University of Tromsø, NO-9037 Tromsø, Norway

(Received 24 September 2003; revised 6 November 2003; accepted 17 November 2003)

Conspicuous sonic click sounds were recorded in the presence of cod (*Gadus morhua*), together with either harp seals (*Pagophilus groenlandicus*), hooded seals (*Cystophora cristata*) or a human diver in a pool. Similar sounds were never recorded in the presence of salmon (*Salmo salar*) together with either seal species, or from either seal or fish species when kept separately in the pool. It is concluded that cod was the source of these sounds and that the clicks were produced only when cod were approached by a swimming predatorlike body. The analyzed click sounds ($n=377$) had the following characteristics (overall averages ± S.D.): peak frequency $=5.95\pm2.22$ kHz; peak-to-peak duration $=0.70\pm0.45$ ms; sound pressure level (received level) $=153.2\pm7.0$ dB re 1 μPa at 1 m. At present the mechanism and purpose of these clicks is not known. However, the circumstances under which they were recorded and some observations on the behavior of the seals both suggest that the clicks could have a predator startling function. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1639106]

PACS numbers: 43.80.Ka, 43.80.n, 43.64.Tk [WA] Pages: 914–919

I. INTRODUCTION

Many fish species are known to produce a variety of sounds, and vocalization among teleost fish has been documented for over 40–50 families, mostly in association with social interactions, and during reproductive periods, in particular (e.g., Brawn, 1961; Myrberg, 1981; Myrberg *et al.*, 1993; Lobel, 1992; Ladich, 1991, 1997; Mann and Lobel, 1998; Nordeide and Kjellsby, 1999). Such fish sounds are usually pulsed, with most of their energy below 3 kHz, and have been characterized as grunts, moans or clicks (e.g., Schneider, 1967; Hawkins and Rasmussen, 1978; Myrberg, 1981). The sounds may be very intense (>130 dB re 1 μPa), which may be important in announcing the size or the physical strength of an individual (Ladich *et al.*, 1992; Myrberg *et al.*, 1993; Crawford *et al.*, 1997). Sound production during nonspawning periods commonly occurs during intraspecific and interspecific aggression, or when fish are disturbed or frightened (e.g., Brawn, 1961; Myrberg, 1981; Ladich, 1997). If sounds are used during agonistic encounters they are usually accompanied with visual agonistic displays (Ladich, 1990; Hawkins, 1993; Mann and Lobel, 1998). Such sounds vary from low frequency grunts and drumming sounds (40–1700 Hz) to higher frequency creaking sounds, clicks and stridulation sounds (1–6 kHz) in various species (Ladich, 1997).

Cod (*Gadus morhua*) have well-developed drumming muscles and are known to produce grunts with their main energy below 1 kHz under a variety of circumstances (McKenzie, 1935; Brawn, 1961; Hawkins and Rasmussen, 1978; Nordeide and Kjellsby, 1999; Midling *et al.*, 2002; Soldal and Totland, 2002). Free-ranging cod have also been found to produce series of knocks with frequencies below 0.6 kHz, in addition to grunts (Midling *et al.*, 2002), but other sounds have, to our knowledge, not been reported for this species.

We recently conducted a bioacoustic study to investigate the potential use of echolocation and acoustical cues by captive harp and hooded seals in connection with hunting of live fish. In that context we recorded click sounds in the presence of both seals and cod that in subsequent follow-up investigations were found to origin from the cod. Here we describe the nature of these cod sounds and the circumstances under which they were recorded.

II. MATERIALS AND METHODS

A. Experimental design

The study was originally designed to identify potential sound production by seals in connection with locating, chasing, and capturing live prey (fish). The present paper only describes those experiments and results that are pertinent for our discovery of some previously undescribed cod sounds. That part of the study involved five types of experiments.

In the first type of (seal-fish) experiments, three sub-adult/adult harp seals (*Pagophilus groenlandicus*) and two sub-adult/adult hooded seals (*Cystophora cristata*; Table I) were maintained, one or two at the time, in a 40 000 liters seawater pool while cod (*Gadus morhua*) were introduced into the pool, 1–3 fish at the time. Any sounds produced in the 20 min following cod introduction (or until all cod had been captured and devoured by the seals) were monitored and recorded for later analyses. These experiments were conducted either with the pool room fully lit (from artificial light sources as well as outdoor light penetrating via eight pool room windows), or with all lights turned off and with blinds on all windows. Light intensity was about 180 lux during light conditions, and 0 lux during dark conditions, as determined with a Grosser Panlux Electronic Luxmeter (Germany). The activities of seals and fish could be surveyed without disturbing the animals during light conditions by use of a video camera (CCD Video, CV252 C, Japan) which was mounted in the ceiling above the pool, and connected to a video screen (CCD, Transvoice Screen, Japan) placed in an

TABLE I. Gender, age, and time in captivity before sound recording experiments were started, for the seals used in the present study.

Animal	Gender	Age	Time in captivity
Harp seals			
G-2	Male	Unknown (subadult–>2 yrs)	8 mo
G-3	Female	Unknown (subadult–>2 yrs)	10 mo
G-5	Female	Unknown (adult–>5 yrs)	5 mo
Hooded seals			
K-1	Female	Unknown (adult–>4 yrs)	4 days
K-66	Male	2.5 yrs	2.5 yrs

adjacent room. A total of 130 cod were introduced, in combinations as shown in Table II.

The second type of (seal-fish) experiments represented a control situation in which juvenile salmon (*Salmo salar*) were introduced instead of cod, while the pool was filled with freshwater instead of seawater. Otherwise these experiments were conducted in exactly the same way as described for the first experimental series. A total of 114 salmon were introduced, in combinations as shown in Table II.

In the third type of experiments, seven cod were introduced into the pool without seals. They were then chased using a hoop net (lights turned on) for 30 min, while any sounds were recorded (Table II).

In the fourth type of experiments, a human diver was placed in the pool while 10 cod were introduced, and sounds were recorded while the diver approached and followed the cod (lights turned on). Altogether 20 cod were confronted with a diver in two 30 min experiments (Table II).

In a fifth type of experiments, cod and salmon (one species at the time) were introduced into the pool (48 cod and 30 salmon, in total), in the abscence of other animals and disturbances, while sounds were recorded for different periods of time (Table II). Recordings of background noise were also conducted in the pool without animals present during a total of 1.5 hours.

B. Experimental animals

The experimental animals were obtained and maintained between experiments as follows: All seals were captured as weanling pups or adults on pack-ice in the Greenland Sea, in connection with scientific expeditions in 1999/2000 under authorization by the Norwegian Government (Royal Ministry of Fisheries). They were brought to the Department of Arctic Biology, University of Tromsø, where they were housed in two 40 000 liters seawater pools under simulated natural light/dark rhythms (corresponding to those at 70° N latitude), and fed a daily ration of fresh-frozen and thawed herring supplemented with a B-vitamin complex, as described by Blix *et al.* (1973). The pool was equipped with a 1.5 m wide wooden ledge along one of its sides which allowed the seals to haul out, and the pool was circulated with seawater at an average temperature of 5 °C, while room air temperature ranged between 8 °C and 10 °C. The age and sex of the seals at the time of the experiments are given in Table I.

A total of 205 cod and 144 salmon were obtained from a marine research station (Kårvika Havbruksstasjon) outside Tromsø, where cod were kept in large net enclosures in the sea, while juvenile salmon were maintained in land-based freshwater tanks, until they were brought to the Department of Arctic Biology for use in the experiments. Upon arrival at the department, they were temporarily maintained in a 500 liters seawater (cod)/freshwater (salmon) tank through which cold (3 °C–5 °C) water was continuously circulated, and oxygenated by pumping air into the water. All cod were juveniles, measuring 20–40 cm in length, except in the experiment with a human diver, in which the cod were somewhat larger, measuring 40–60 cm, while the juvenile salmon measured 20–30 cm in length (Table II). All fish that were transported to the facilities were used in experiments within 2–3 days; hence, feeding of the fish was not needed. All experiments were approved by the Norwegian National Animal Research Authority (NARA).

TABLE II. The number of hours of sound recordings, numbers and size of fish used, and numbers of recorded clicks, with mean peak frequency, under various experimental conditions.

Condition	Recording time (h:min)	Number of fishes	Fish size (cm)	Clicks recorded			F_{peak} (kHz)
				Light	Dark	Total	
Cod+harps	13:30	65	20–40	54	83	137	5.99±1.82
Cod+hoods	8:30	65	20–40	88	79	167	6.66±2.48
Salmon+harps	4:50	38	20–30	0	0	0	
Salmon+hoods	16:30	76	20–30	0	0	0	
Cod+hoop net	0:30	7	20–40	0		0	
Cod+human	1:00	20	40–60	73		73	4.26±1.07[a]
Cod	2:00	48	20–40	0		0	
Salmon	1:00	30	20–30	0		0	

[a]F_{peak} significantly different from those recorded in the presence of harp or hooded seals [$p<0.001$, Mann-Whitney $U=1995.5$ (harps) and 2178.0 (hoods)].

C. Sound recordings

The 40 000 liters experimentation pool (measuring 5.8 m in length by 5.6 m in width and with water depth ranging between 1.2 and 1.4 m) was made out of plywood, covered with fiberglass and a polyester-based topcoat. All pool machinery (pumps supplying new and recirculated water and water filtering machines) was turned off during experiments. An omnidirectional Brüel and Kjær Type 8103 hydrophone [frequency range 0.1–180 kHz flat (± 2 dB) up to 100 kHz, and sensitivity -211.7 ± 0.25 dB re 1 V μPa^{-1}; Brüel and Kjær, Denmark] was suspended from the ceiling of the pool room into the center of the pool at a water depth of approximately 65 cm (i.e., midway between bottom and surface). The hydrophone and cable were protected by a wire mesh (wire, $\varnothing = 1$ mm) to avoid mauling by the seals. The hydrophone cable was connected to a Nexus Type 2692 amplifier (Brüel and Kjær, Denmark) in the adjacent surveillance room, which in turn sent signals to a loudspeaker, an oscilloscope (Type HM205-3 Hameg, Germany), and a Racal Store 4DS analog instrumentation tape recorder (Hardley Industrial Estate, England), respectively. The recorder was operated at a tape speed of 30 inches per second, enabling flat recordings from 0.2 to 150 kHz (± 3 dB), with a signal-to-noise ratio of 40 dB. It was calibrated using a Wavetek sound generator (Wavetek 182, Wavetek Corp., IN, USA).

Sound recordings were started in connection with fish introduction and were continued until all fish had been devoured (in the seal-fish experiments) or until the experiment was terminated 20 min after the introduction of fish. Live fish were kept in a dark room for at least 30 min prior to the darkness experiments, assuming that this would provide the fish with sufficient time to adapt to dark conditions.

D. Sound analyses

All recorded sounds of possible animal origin were identified from the recorded tapes by replaying them at various tape speeds, both via a loudspeaker and an oscilloscope. Identified sounds were then digitalized and analyzed on a personal computer with a sound card, using analysis software programs (Bat Sound Pro, Petterson Electronics 1996, Sweden, or MATLAB, The MathWorks Inc. 1984–1999, MA, USA).

Minimum (F_{min}), maximum (F_{max}), and peak (F_{peak}) frequencies were determined, either using the energy distribution area of the spectrograms, or by using a power spectrum analysis. Short-duration sounds were also described by their peak-to-peak duration (t_{peak}), centroid frequency (F_{cent}), and the root-mean-square bandwidth (BW_{rms}). These measurements were made using a program constructed on MATLAB. Received sound pressure levels (SPL, in dB re 1 μPa) were estimated based on the voltage output (as determined by use of the oscilloscope), taking the hydrophone and amplifier sensitivities into account.

E. Statistics

Sound characteristics are presented as arithmetic means with standard deviations, or as medians. Between-experiment comparisons of sound characteristics were made using nonparametric tests (Kruskal-Wallis and Mann-Whitney), using SPSS for Windows, version 11.0 (SPSS Inc., IL, USA), with differences considered significant at a level of $p < 0.05$.

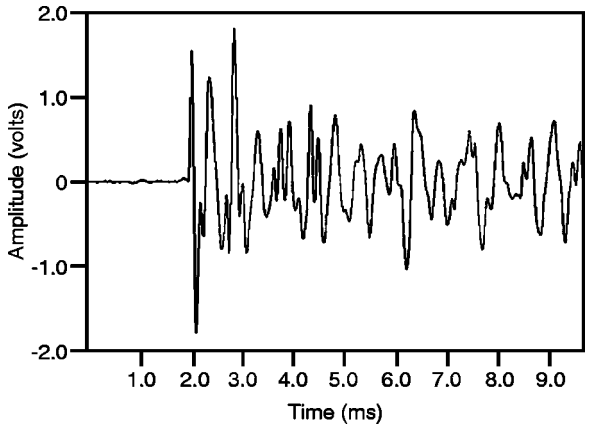


FIG. 1. Short waveform of a single click recorded in the presence of cod and a male harp seal. Duration (t_{peak}), ca. 0.3 ms.

III. RESULTS

A total of 377 sonic clicks of similar characteristics were analyzed during the course of the study. The clicks either occurred as singular clicks ($n = 163$, Figs. 1 and 2), clicks preceded by a pre-click ($n = 80$, Fig. 3) or double/triplet

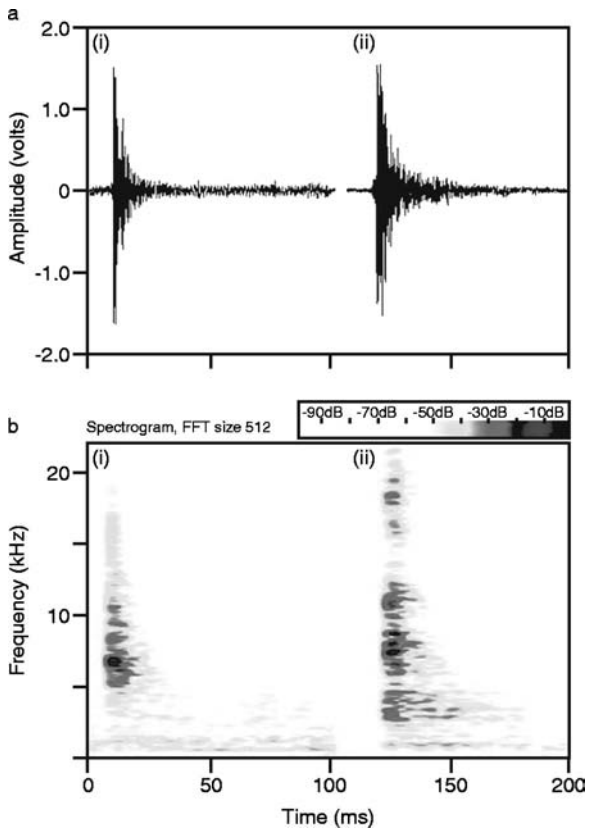


FIG. 2. (a) Oscillogram of a single click recorded in the presence of cod and (i) a human diver (SPL, 153 dB re 1 μPa at 1 m), and (ii) a male harp seal (SPL, 154 dB re 1 μPa at 1 m). (b) Spectrogram (Hanning window) of a single click recorded in the presence of cod and (i) a human diver (frequency range, 4.6–11.2 kHz; F_{peak}, 6.7 kHz), and (ii) a male harp seal (frequency range, 2.7–18.7 kHz; F_{peak}, 7.6 kHz).

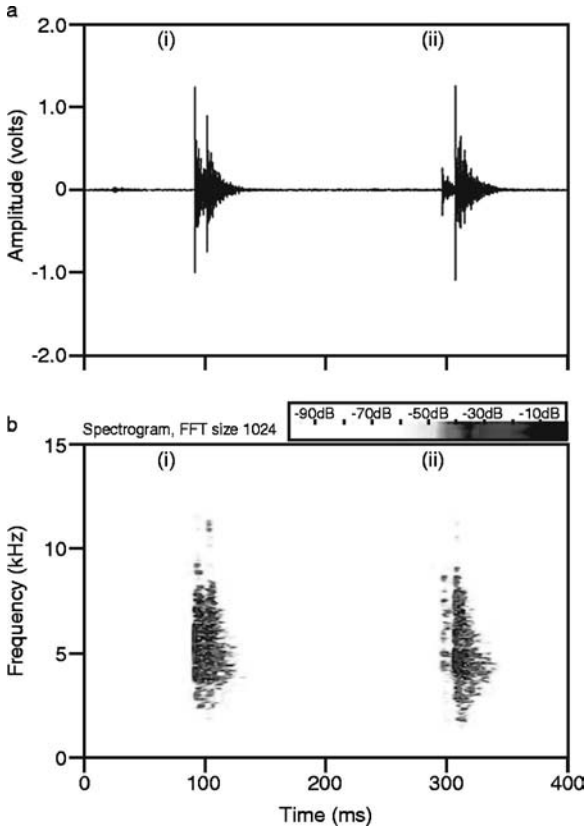


FIG. 3. (a) Oscillogram of (i) a double click (SPL of clicks, 152 dB/150 dB re 1 μPa at 1 m; interval between clicks, 19.2 ms) and (ii) a click with a preclick (SPL of preclick and click, 142 and 152 dB re 1 μPa at 1 m, respectively; interval between preclick and click, 22 ms), as recorded in the presence of cod and a male harp seal. Durations (t_{peak}), 0.3–2 ms. (b) Spectrogram (Hanning window) of a double click (i), and a click with a preclick (ii), as recorded in the presence of cod and a male harp seal. Frequency range, 3.6–8.3 kHz; F_{peak}, ca. 5 kHz.

clicks ($n=134$, Fig. 3), all having a very sudden onset and similar frequency characteristics (Table III). The overall average F_{cent} and BW_{rms} were 7.06 ± 3.79 kHz and 4.15 ± 2.77 kHz, respectively, while the average SPL was estimated to be 153.2 ± 7.0 dB re 1 μPa at 1 m, ranging between 139.2 ± 7.0 and 163.2 ± 7.0 dB re 1 μPa at distances of 0.2 m and 3 m, respectively (based on $n=184$ analyzed clicks).

In the type 1 experiments (cod with seals), 304 analyzable clicks were recorded, distributed with $n=167$ clicks in the presence of cod and hooded seals, and $n=137$ clicks in the presence of cod and harp seals (Table II). In contrast, no similar clicks were recorded in any of the type 2 experiments (salmon with seals), in the type 3 experiments with cod and a hoop net, or in the control experiments (type 5) with either cod or salmon alone in the pool. However, when cod were introduced into the pool with a human diver (type 4 experiments), a total of 73 analyzable clicks were recorded. The characteristics of clicks recorded under various experimental conditions differed in that clicks recorded in the presence of a human diver were of significantly lower frequencies than those recorded in the presence of seals [$p<0.001$, Mann-Whitney $U=1995.5$ (harp seals) and 2178.0 (hooded seals); Table II]. However, the average F_{peak} of clicks recorded in the presence of cod and harp seals did not differ significantly from that of clicks recorded in the presence of cod and hooded seals ($p=0.075$, Mann-Whitney $U=10\,086.0$). Cod showed a strong avoidance behavior both when approached by the human diver and by the seals.

In addition to clicks, 2 grunts were also recorded when cod were confronted with a human diver. These grunts had a t_{peak} of 190 ms and F_{peak} between 0.50 and 0.73 kHz. The average SPL was estimated to be 133 ± 3.0 dB re 1 μPa at 1 m, ranging between 116 and 146 dB re 1 μPa, at distances of 0.2 and 3 m, respectively.

IV. DISCUSSION

Sound recordings in a small pool are associated with obvious problems, primarily caused by reverberations. During the present study echoes disturbed signals by often overlapping the second signal. Also, the reported clicks were unexpectedly loud, and some signals therefore tended to overload the sound recording system. Both these factors made analyses of sounds sometimes difficult and reduced the number of clicks being successfully analyzed. We were nevertheless able to record and analyze conspicuous and characteristic sonic clicks in the presence of cod and seals (both harp seals and hooded seals), and in the presence of cod and a human diver. However, we were unable to record similar clicks in the presence of salmon and seals, or while salmon or cod were alone in the pool, either undisturbed or when being chased with a hoop net (cod). Also, clicks were not recorded while seals were alone in the pool, or while no animals were present at all. These findings strongly suggest (1) that the recorded clicks originated from cod and (2) that the sounds were produced by cod when approached by a predator/predatorlike body. To our knowledge, this is the first published documentation of the production of this type of click sounds by cod.

Cod have previously been documented to produce mainly low-frequency sounds under a variety of circumstances, grunts being particularly common in connection with aggression or when the fish is frightened, during, as

TABLE III. Frequency characteristics of different categories of clicks recorded in the presence of cod.

Click type	Number	F_{peak} (kHz)	F_{min} (kHz)	F_{max} (kHz)	F_{cent} (kHz)	BW_{rms} (kHz)	t_{peak} (ms)	interval median (ms)
Single clicks	163	6.17 ± 2.11	3.57 ± 1.92	11.70 ± 4.47	7.38 ± 3.72 ($n=83$)	4.16 ± 2.55 ($n=83$)	0.81 ± 0.51 ($n=148$)	
Preclicks	40	6.60 ± 2.66	3.26 ± 2.24	11.92 ± 5.78	8.55 ± 5.65 ($n=8$)	4.41 ± 2.94 ($n=8$)	0.45 ± 0.39 ($n=39$)	24
Clicks after preclick	40	5.72 ± 2.47	2.69 ± 2.04	11.14 ± 5.11	8.05 ± 3.03 ($n=8$)	3.51 ± 2.13 ($n=8$)	0.60 ± 0.40 ($n=32$)	
Double/triplets	134	5.55 ± 2.07	3.07 ± 2.12	10.44 ± 6.47	5.35 ± 1.26 ($n=10$)	1.44 ± 0.28 ($n=10$)	0.67 ± 0.34 ($n=112$)	22
Total	377	5.95 ± 2.22	3.26 ± 2.05	11.23 ± 5.48	7.06 ± 3.79 ($n=109$)	4.15 ± 2.77 ($n=109$)	0.70 ± 0.45 ($n=331$)	

well as outside, the spawning season (e.g., Brawn, 1961; Hawkins and Rasmussen, 1978; Soldal and Totland, 2002). Such grunts have their main energy below 1 kHz, durations of 60–200 ms and source levels of 120–133 dB re 1 μPa at 1 m (Brawn, 1961; Hawkins and Rasmussen, 1978; Midling *et al.*, 2002; Nordeide and Kjellsby, 1999), thus, being similar to the two grunts recorded in the present study. It should be noted, in this context, that the recording equipment was not set to record sounds below 200 Hz and that background low-frequency noise caused by swimming movements of the animals may have disguised some lower-frequency grunts ($<$500 Hz) in the present study. Moreover, Brawn (1961) reported that only cod with body lengths of $>$37 cm produced grunt sounds, and the fact that most cod used in the present study were smaller than this might contribute to explain the sparse occurrence of grunts. Series of low-frequency knocks (below 0.6 kHz) have also been recorded from free-ranging cod (Midling *et al.*, 2002), but no report on medium-frequency clicks similar to those reported in the present study seems to exist. This may be related to the fact that the clicks apparently were produced only when cod were confronted with, and approached by, a predatorlike body, like a seal or a human diver, while no clicks were recorded when cod were chased with a hoop net. Brawn (1961) reported that cod that were confronted with a conger eel (*Conger oceanicus*) produced low-frequency grunts with a peak frequency of about 50 Hz, but no mention was made of any click sounds. However, the production of intense click sounds is known from other species of fish, such as *Pollimyrus isidori* during social interaction (Crawford *et al.*, 1997), *Botia horae* during juvenile agonistic encounters (Valinsky and Rigley, 1981), and during the defense of territory in *Lagodon rhomboides* (Caldwell and Caldwell, 1967).

We have no data to suggest a possible mechanism for click production in cod, but the frequency range of clicks ($>$1.5 kHz) suggests that it represents some form of stridulation (Kaatz, 2002), rather than resulting from the contraction of the striated drumming muscles surrounding the swimbladder wall, which is the known mechanism by which cod produce grunts (Brawn, 1961). It may be speculated that the stridulation mechanism involves a sudden movement of skeletal parts against each other, possibly as the cod makes an escape movement when a predator approaches the fish. In fact, Brawn (1961) suggested that sudden movements accomplished sound production in cod. Yet another possibility is the production of distinct pressure pulses generated by rapid swim movements, as reported in schooling fish by Gray and Denton (1991). However, the sound was absent when the cod swam away from the hoop net, and, thus, appears to be actively produced only in particularly threatening situations. In contrast, salmon apparently produced no sounds while escaping the seals, and this also seems to have been the case with regard to Antarctic cod (*Dissostichus mawsoni*), which were reported to remain silent upon escape from a Weddell seal, *Leptonychotes weddellii* (Davis *et al.*, 1999). It should be noted, also, that no click or other sound was recorded in the present study after the cod had been seized by the seal, which is in accordance with observations made by Brawn (1961). The click sounds that were recorded from cod in the presence of a human diver were somewhat, but nevertheless significantly, lower in frequency than those recorded in the presence of seals. This may be related to the fact that the cod used in the former experiments were larger in size than those used in the latter (Table II), for reasons beyond our control. This hypothesis presumes that the frequency of produced sounds is size-related, as shown to be the case for other sound producing mechanisms in fish (e.g., Connaughton *et al.*, 2002).

The function of the clicks is also open for speculation. The fact that they only occurred in connection with the approach of a predatorlike body is suggestive of some warning or protective function. The use of intense sounds in similar situations is well known from studies of other fish species (e.g., after disturbance by a competitor or a predator; for review see Myrberg, 1981; Ladich, 1997; Mann and Lobel, 1998), including cod (Brawn, 1961). The presently reported click sounds are outside the optimal frequency range for auditory sensitivity reported for cod (60–310 Hz, Chapman and Hawkins, 1973), making it less likely that they act as warning signals for conspecifics, but audiogram tests have not been conducted at frequencies above 1 kHz. Moreover, Astrup and Møhl (1993) reported that cod are able to detect ultrasound (38 kHz), at least at high intensities (194.4 dB re 1 μPa). It therefore cannot be excluded that the presently reported click sounds could be detected by the cod.

An alternative option, however, would be that the sounds are used as a predator deterrent or startling mechanism, along the lines suggested by Mahajan (1963), Pfeiffer and Eisenberg (1965), and Myrberg (1981), and as demonstrated for some insects in their interaction with echolocating predator bats (e.g., Møhl and Miller, 1976; Miller and Surlykke, 2001). The cod clicks had their main energy within the best hearing range of harp seals (Terhune and Ronald, 1972) and of several other marine mammals as well (see review by Richardson *et al.*, 1996), and were therefore probably readily detected by the seals. Circumstantial evidence for a startling/deterrent effect were actually obtained in the present study: Even though the hooded seals appeared unaffected by the clicks, the harp seals were, in fact, observed to abort their approach towards clicking cod on several occasions and also captured a significantly lower proportion of (clicking) cod compared to the similar sized, though non-clicking, salmon (Vester *et al.*, 2001; Vester, 2003). This observation might explain why cod, which are very common in the Barents Sea where their distribution overlaps with that of harp seals, are not found in any abundance in harp seal stomachs, which are instead rich in capelin, herring, and invertebrates (e.g., Nilssen, 1995) that are also readily available in the area.

In conclusion, the present study has demonstrated that cod are capable of producing intense sonic clicks not previously described. The sound production only occurred when cod were confronted with large predatorlike bodies (seals and a human diver) and may act as a predator deterrent.

ACKNOWLEDGMENTS

We are grateful to Rolf Rødven for his kind assistance in connection with the experiments with a human diver, and to

Magnus Wahlberg at Department of Zoophysiology, Institute of Biological Sciences, University of Aarhus, for help and advice in connection with the sound analyses. This study was financed in part by contributions from the Roald Amundsen Center for Arctic Research, University of Tromsø.

Astrup, J., and Møhl, B. (**1993**). ''Detection of intense ultrasound by the cod *Gadus morhua*,'' J. Exp. Biol. **182**, 71–80.

Blix, A. S., Iversen, J., and Paasche, A. (**1973**). ''On the feeding of young hooded seals (*Cystophora cristata*) and harp seals (*Pagophilus groenlandicus*) in captivity,'' Norw. J. Zool. **21**, 55–58.

Brawn, V. M. (**1961**). ''Sound production by the cod (*Gadus morhua* L.),'' Behaviour **18**, 239–255.

Caldwell, D. K., and Caldwell, M. C. (**1967**). ''Underwater sounds associated with aggressive behaviour in defence of territory by the pinfish, *Lagodon rhomboides*,'' Bull. South. Calif. Acad. Sci. **66**, 69–75.

Chapman, C. J., and Hawkins, A. D. (**1973**). ''A field study of hearing in the cod, *Gadus morhua*,'' J. Comp. Physiol. **85**, 147–167.

Connaughton, M. A., Fine, M. L., and Taylor, M. H. (**2002**). ''Weakfish sonic muscle: influence of size, temperature and season,'' J. Exp. Biol. **205**, 2183–2188.

Crawford, J. D., Jacob, P., and Bénech, V. (**1997**). ''Sound production and reproductive ecology of strongly acoustic fish in Africa: *Pollimyrus isidori*, Mormyridae,'' Behaviour **134**, 677–725.

Davis, R. W., Fuiman, L. A., Williams, T. M., Collier, S. O., Hagey, W. P., Kanatous, S. B., Kohin, S., and Horning, M. (**1999**). ''Hunting behavior of a marine mammal beneath the Antarctic fast ice,'' Science **283**, 993–996.

Gray, J. A. B., and Denton, E. J. (**1991**). ''Fast pressure pulses and communication between fish,'' J. Mar. Biol. Assoc. (U.K.) **71**, 83–106.

Hawkins, A. D. (**1993**). ''Underwater sound and fish behaviour,'' in *Behaviour of Teleost Fishes*, 2nd ed., edited by T. J. Pitcher (Chapman and Hall, Dordrecht), pp. 129–169.

Hawkins, A. D., and Rasmussen, K. J. (**1978**). ''The calls of gadoid fish,'' J. Mar. Biol. Assoc. (U.K.) **58**, 891–911.

Kaatz, I. M. (**2002**). ''Multiple sound-producing mechanisms in teleost fishes and hypotheses regarding their behavioural significance,'' Bioacoustics **12**, 230–233.

Ladich, F. (**1990**). ''Vocalisation during agonistic behaviour in *Cottus gobio* (Cottidae): An acoustic threat display,'' Ethology **84**, 193–201.

Ladich, F. (**1991**). ''Fische schweigen nicht—Lautbildung, Hoeren and akustische Kommunikation bei Fischen,'' (''Fish are not silent—sound production, hearing and acoustic communication in fishes''), Naturwiss. Rundsch. **10**, 379–384.

Ladich, F., Brittinger, W., and Kratochvil, H. (**1992**). ''Significance of agonistic vocalisation in the croaking gourami (*Trichopsis vittatus*, Teleostei),'' Ethology **90**, 307–314.

Ladich, F. (**1997**). ''Agonistic behaviour and significance of sound in vocalizing fish,'' Mar. Fresh. Behav. Physiol. **29**, 1285–1288.

Lobel, P. S. (**1992**). ''Sound production by spawning fishes,'' Environ. Biol. Fish. **33**, 351–358.

Mahajan, C. L. (**1963**). ''Sound producing apparatus in an Indian catfish *Sisor rhabdophorus* Hamilton,'' J. Linn. Soc. Lond. Zool. **44**, 721–724.

Mann, D. A., and Lobel, P. S. (**1998**). ''Acoustic behavior of the damselfish *Dascyllus albisella*: behavioral and geographic variation,'' Environ. Biol. Fish. **51**, 421–428.

McKenzie, R. A. (**1935**). ''Codfish in captivity,'' Prog. Rep. Atl. Biol. Sta. **16**, 7–10.

Midling, K., Soldal, A. V., Fosseidengen, J. E., and Øvredal, J. T. (**2002**). ''Calls of the Atlantic cod: Does captivity restrict their vocal repertoire?,'' Bioacoustics **12**, 233–235.

Miller, L. A., and Surlykke, A. (**2001**). ''How some insects detect and avoid being eaten by bats: Tactics and countertactics of prey and predator,'' BioScience **51**, 570–581.

Myberg, Jr., A. A. (**1981**). ''Sound communication and interception in fishes,'' in *Hearing and Sound Communication in Fishes*, edited by W. N. Tavolga, A. N. Popper, and R. R. Fay (Springer, New York), pp. 395–426.

Myrberg, Jr., A. A., Ha, S. J., and Shamblott, M. J. (**1993**). ''The sound of bicolor damselfish (*Pomacentrus partitus*): predictors of body size and spectral basis for individual recognition and assessment,'' J. Acoust. Soc. Am. **94**, 3067–3070.

Møhl, B., and Miller, L. A. (**1976**). ''Ultrasonic clicks produced by the peacock butterfly: A possible bat-repellent mechanism,'' J. Exp. Biol. **64**, 639–644.

Nilssen, K. T. (**1995**). ''Seasonal distribution, condition and feeding habits of Barents Sea harp seals (*Phoca groenlandica*),'' in *Whales, Seals, Fish and Man*, edited by A. S. Blix, L. Walløe, and Ø. Ulltang (Elsevier, Amsterdam), pp. 241–254.

Nordeide, J. T., and Kjellsby, E. (**1999**). ''Sound from spawning cod at their spawning grounds,'' ICES J. Mar. Sci. **56**, 326–332.

Pfeiffer, W., and Eisenberg, J. F. (**1965**). ''Die Lauterzeugung der Dornwelse (Doradidae) und der Fiederbartwelse (Mochokidae),'' [''Sound production of thorny catfishes (Doradidae) and squeakers (Mochokidae)''], Z. Morphol. Oekol. Tiere **54**, 669–679.

Richardson, W. J., Greene, Jr., C. R., Malme, C. I., and Thomson, D. H. (**1996**). *Marine Mammals and Noise* (Academic, San Diego).

Schneider, H. (**1967**). ''Morphology and physiology of sound-producing mechanisms in teleost fishes,'' in *Marine Bio-acoustics*, edited by W. N. Tavolga (Pergamon, Oxford, England), Vol. 2, pp. 135–158.

Soldal, A. V., and Totland, B. (**2002**). ''Biologisk lyd brukt til selektivt fiske,'' (''Biological sounds used for selective fishing''), Audio recordings on compact disc. Marine Research Institute, Bergen, Norway.

Terhune, J. M., and Ronald, K. (**1972**). ''The harp seal, *Pagophilus groenlandicus* (Erxleben, 1777). III. The underwater audiogram,'' Can. J. Zool. **50**, 565–569.

Valinsky, W., and Rigley, L. (**1981**). ''Function of sound production by the skunk loach *Botia horae* (Pisces, Cobitidae),'' Z. Tierpsychol. **55**, 161–172.

Vester, H. I. (**2003**). ''Do Arctic seals use echolocation?,'' M.S. thesis, Department of Arctic Biology, University of Tromsø, Norway.

Vester, H. I., Folkow, L. P., and Blix, A. S. (**2001**). ''Click sounds produced by cod (*Gadus morhua*)—a possible anti-predator behaviour?,'' 17th International Congress on Acoustics, Rome, September 2–7, 2001. Rivista Italiana di Acustica (Special Issue), **25** (1), Abstract 4P.42

An echolocation model for range discrimination of multiple closely spaced objects: Transformation of spectrogram into the reflected intensity distribution

Ikuo Matsuo,[a] Kenji Kunugiyama, and Masafumi Yano
Research Institute of Electrical Communication, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai, 980-8577, Japan

(Received 29 July 2003; revised 15 November 2003; accepted 24 November 2003)

Using frequency-modulated echolocation, bats can discriminate the range of objects with an accuracy of less than a millimeter. However, bats' echolocation mechanism is not well understood. The delay separation of three or more closely spaced objects can be determined through analysis of the echo spectrum. However, delay times cannot be properly correlated with objects using only the echo spectrum because the sequence of delay separations cannot be determined without information on temporal changes in the interference pattern of the echoes. To illustrate this, Gaussian chirplets with a carrier frequency compatible with bat emission sweep rates were used. The delay time for object 1, T_1, can be estimated from the echo spectrum around the onset time. The delay time for object 2 is obtained by adding T_1 to the delay separation between objects 1 and 2 (extracted from the first appearance of interference effects). Further objects can be located in sequence by this same procedure. This model can determine delay times for three or more closely spaced objects with an accuracy of about 1 μs, when all the objects are located within 30 μs of delay separation. This model is applicable for the range discrimination of objects having different reflected intensities and in a noisy environment (0-dB signal-to-noise ratio) while the cross-correlation method is hard to apply to these problems. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1642626]

PACS numbers: 43.80.Lb, 43.64.Bt, 43.80.Jz [WA] Pages: 920–928

I. INTRODUCTION

Bats emit trains of high-frequency sound waves and can locate an individual object among three or more objects using the temporally overlapping echoes of the emitted waves. (Griffin, 1958; Simmons *et al.*, 1995). Experimental evidence indicates that FM bats are capable of locating an object with a submicrosecond accuracy (Menne *et al.*, 1989; Moss and Schnitzler, 1989; Simmons, 1979; Simmons *et al.*, 1990). For two or more closely spaced objects, the echoes will interfere with one another. Convolution of the constant frequency (CF) carrier wave with the echoes allows delay separations among them to be estimated from the spectrum or spectrogram of echoes (Matsuo *et al.*, 2001; Saillant *et al.*, 1993; Neretti *et al.*, 2003). It is, however, difficult to determine the temporal sequence of delay separations since this transformation cannot extract temporal changes in the echo spectrum.

To accurately locate three or more objects, it is necessary to extract such temporal changes in the echo spectrum. First, it is necessary to accurately determine the delay time T_1 for the first object, because the error in T_1 seriously affects the determination of the temporal sequence of delay separations. Although it has been found that bats can measure each delay time T_i for two closely spaced objects with microsecond accuracy (Simmons *et al.*, 1995, 1996, 1998), models to accurately determine T_1 have not yet been proposed. T_1 can be approximated using the cross-correlation function of the spectra of emitted and reflected waves (Matsuo *et al.*, 2001). However, it is difficult to accurately determine T_1 from this cross-correlation function because the integration time of cochlear filters is large (several hundreds of microseconds). For the accurate determination of T_1, a new method is proposed which utilizes the echo spectrum. The second requirement for the determination of the temporal sequence of delay separations is the extraction of temporal changes in the interference patterns from the echo spectra. It is difficult to extract these temporal changes through the application of the usual method of convolving the echoes with the constant frequency (CF) carrier wave to transform echoes into the echo spectra, because this method is relatively insensitive to temporal changes. In order to solve these problems, we introduce Gaussian chirplets for which the carrier frequency agrees with the sweep rate of emission (Mann and Haykin, 1992, 1995), and present a new method for the accurate determination of each delay time T_i for three or more closely spaced objects.

II. MODEL

A. Characteristics of emission and echo

1. Emission

Our model uses FM emission, because the purpose of the model is to identify an object's location. A single emission event from an echolocating bat consists of a principal frequency that sweeps through a range of values (usually from about 10 kHz to about 100 to 200 kHz), together with single or multiple harmonics (Fenton and Bell, 1981). Emissions for *Myotis lucifugus* have one type of fundamental FM emission. A single emission is of duration 2 to 5 ms, and

[a]Electronic mail: matsuo@riec.tohoku.ac.jp

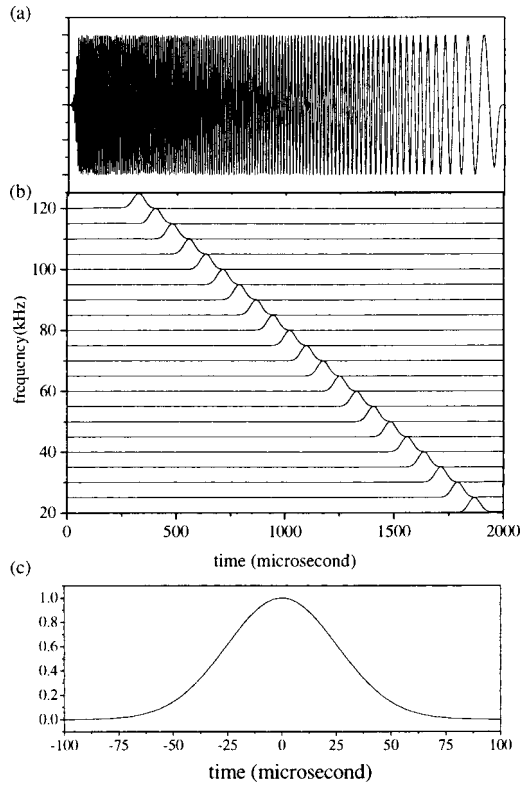


FIG. 1. (a) Emission waveform. (b) Emission spectrogram. (c) Temporal emission pattern W_{emi} corresponding to the spectrogram in one bandpass filter (with center frequency 70 kHz).

sweeps frequencies from 100 kHz down to 40 kHz (Fenton and Bell, 1981). This model uses an emission with a single harmonic, linear FM sweep. The emission spectrum used has constant amplitude from 20 to 120 kHz. For this reason, the 2-ms-long FM emission begins at 135 kHz and sweeps down to 5 kHz. The rise and fall times are 50 μs. Figure 1 shows the emitted waveform.

2. Echo

The following conditions were assumed to simplify the simulation:

(i) Each object is a point object. An extended object is defined as consisting of a distribution of acoustically reflecting points, called ''glints.'' (Altes, 1976).
(ii) Reflection does not modify the phase of the sound wave.

The plot of reflected intensity vs distance (equivalent to delay time) is called the ''reflected intensity distribution.''

B. Transformation of the waveform into the spectrogram by Gaussian chirplets

Waveforms for emission and echoes from many closely spaced objects were put into this model. These waveforms were transformed into spectrograms in a manner that simulates this process in the mammalian cochlea. The carrier wave previously used for cochlear bandpass filters was a wave of a constant frequency (CF) corresponding to the center frequency of the bandpass filter (Matsuo *et al.*, 2001; Neretti *et al.*, 2003; Saillant *et al.*, 1993). It was previously impossible to extract temporal changes in the interference patterns of echoes due to the convolution of the CF carrier wave with both the FM wave emitted and the echoes. In order to extract temporal changes in the interference pattern, we used a Gaussian chirplet with a carrier frequency in agreement with the sweep rate of emission (Mann and Haykin, 1992, 1995). The temporal characteristics of such a filter can be described by

$$F(f_j,t)=\exp\left(-\frac{t^2}{\alpha}\right)\exp\left(2\pi j\left(f_j t+\frac{1}{2}st^2\right)\right), \tag{1}$$

where f_j (kHz) is the center frequency for the jth bandpass filter, t is time (s), s is the sweep rate of the emitted wave, and α is a parameter for the width of the window function. This bandpass filter bank is composed of 101 filters. The center frequencies of these filters range from 20 to 120 kHz, positioned at regular frequency intervals. At 50% of peak height, the width of the window function for all the filters was 160 μs, with $Q_{10\,\text{dB}}$ values ranging from 1 at 20 kHz to 6 at 120 kHz. The waveforms for both the emitted waves and echoes were transformed into the spectrogram $P(f,t)$ by convolution with this filter in Eq. (1). Figure 1(b) shows the outputs of the cochlear filters for the emitted waveform. Figure 1(c) shows the temporal pattern corresponding to the spectrogram $P(f,t)$ of emission in one filter (with center frequency 70 kHz). The shapes of the temporal patterns corresponding to the spectrogram $P(f,t)$ in all bandpass filters are the same due to use of the linear FM emission. This temporal pattern is termed ''W_{emi}.''

In order to demonstrate the output from cochlear filters for echoes originating from many closely spaced objects, we considered the situation of three such objects. The delay times for these objects were 4100, 4150, and 4225 μs and the reflectivity of each object was 1. Figures 2(a) and (b) display the waveforms of the echoes and the output of the Gaussian chirplets, respectively. This spectrogram $P(f,t)$ was transformed into a range-frequency pattern $S_{\text{echo}}(f,\tau)$, with 10-μs interval by compensating for the sweep rate of the emission frequency modulation. Since the start time for emission was 0, this compensating time was represented by τ, which implies the range corresponding to the delay time. The solid line in Fig. 2(c) represents the range–frequency pattern computed using Gaussian chirplets, and the dotted line represents the range–frequency pattern computed using bandpass filters of the CF carrier wave with the same window function as the Gaussian chirplet. Figure 2(c) shows that temporal changes in interference patterns can be extracted by convolution of the Gaussian chirplet.

C. Correspondence of the output of cochlear filters with the reflected intensity distribution along the range axis

In order to demonstrate the correspondence of the outputs of cochlear filters with the reflected intensity distribution along the range axis (T), range-frequency patterns $S_{\text{ref}}(f,\tau)$ were theoretically computed from the reflected intensity distribution in Fig. 3(a). This pattern, computed by multiplying the reflected intensity distribution by values of W_{emi} whose peak time corresponds to τ [Fig. 3(a)], is called

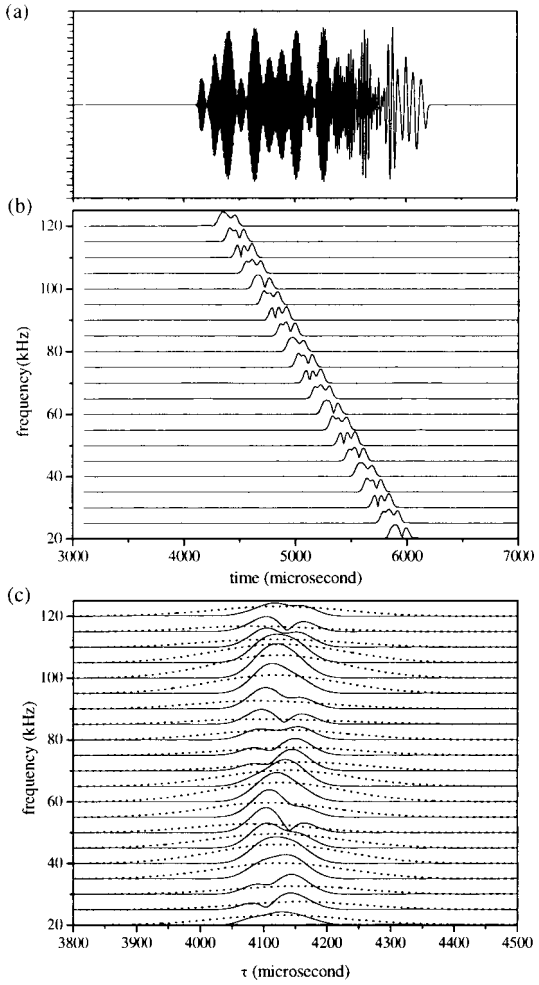


FIG. 2. (a) Echo waveform. T_i for the three objects are 4100, 4150, and 4225 μs and the reflectivities are 1. (b) Echo spectrogram. (c) Range–frequency pattern computed from the spectrogram. The solid and dotted curves represent the results computed by convolution of Gaussian chirplets and those computed by convolution of the CF carrier wave, respectively.

the "windowed reflected intensity distribution," $R(T,\tau)$. Figure 3(b) shows the windowed reflected intensity distribution R that is represented along the range axis at each delay τ. Values of R at T_i are changed dependent on the delay τ because of the narrow width of W_{emi}. $S_{\text{ref}}(f,\tau)$ was computed by Fourier transforming the windowed reflected intensity distribution $R(T,\tau)$. As shown in Figs. 3(c) and (d), the spectral shape of the $S_{\text{echo}}(f,\tau)$ closely corresponds to that of the $S_{\text{ref}}(f,\tau)$ at each τ.

D. The determination of delay times for closely spaced objects

Determination of each delay time for three or more closely spaced objects requires estimation of T_1 and the temporal sequence of delay separations from the range–frequency pattern. T_1 was determined from the echo spectra around the onset delay τ_{on} since here, these are mainly influenced by the echo of the first object. By considering the situation of a single object, it is possible to clarify the relationship between S_{ref} and the delay time. For example, in Fig. 4(a), the delay time for the object is 4100 μs. Figures 4(b) and (c) show the windowed reflected intensity distribution is computed by multiplying the intensity distribution of this reflection by emission values W_{emi}, whose peak times are 4050 and 4060 μs, respectively. At $\tau=4050$ and 4060

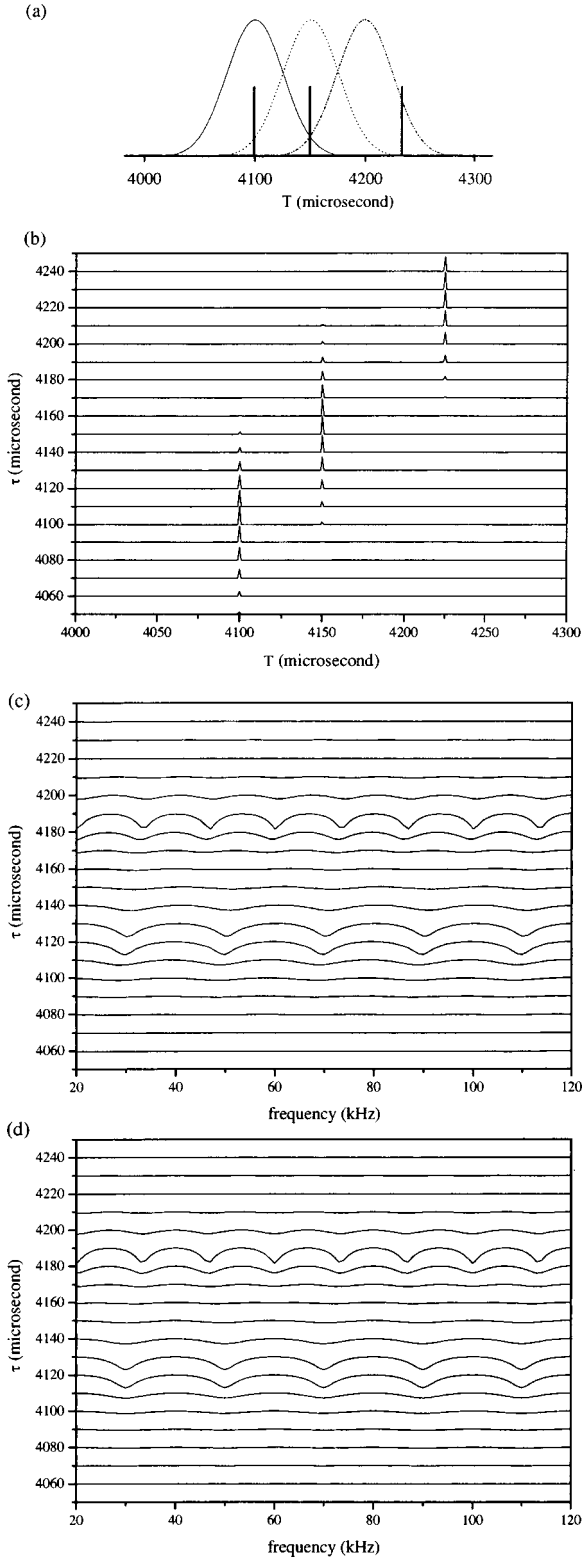


FIG. 3. Correspondence of the output from cochlear filters with the reflected intensity distribution. (a) Reflected intensity distribution and W_{emi}. The peak times of W_{emi} correspond to 4100, 4150, 4200 μs, respectively. (b) Windowed reflected intensity distribution, R. (c) Range–frequency pattern S_{echo} computed by convolution with Gaussian chirplets. (d) Range–frequency pattern S_{ref} computed from the reflected intensity distribution.

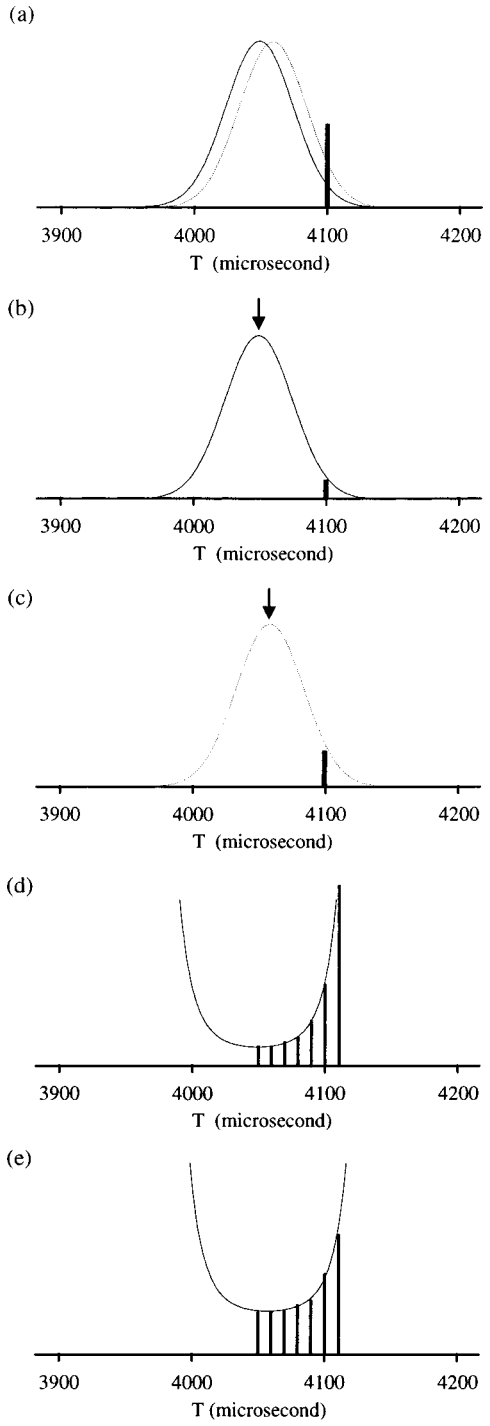


FIG. 4. Determination of T_1. (a) Intensity distribution of a single reflection and the W_{emi} whose peak times correspond to 4050 and 4060 μs, respectively. (b) Windowed reflected intensity distribution at $\tau=4050$ μs. The arrow shows the 4050-μs location. (c) Windowed reflected intensity distribution at $\tau=4060$ μs. The arrow shows the 4060-μs location. (d) Candidates for the reflected intensity distribution are estimated from S_{echo} at the onset delay time τ_{on}. (e) Candidates for the reflected intensity distribution are estimated from S_{echo} at 10 μs after τ_{on}. T_1 is determined by estimating the correspondence between the reflectivity of candidates at each delay time.

μs, S_{ref} has a constant amplitude, but the average of S_{ref} for these two times is not equal. T_1 and the corresponding reflectivity r_1 are uniquely determined from the averages of the two spectra at the onset delay time τ_{on} and 10 μs later, because the average of S_{ref} at an arbitrary τ is dependent upon both T_1 and r_1. As shown in Fig. 4(d), candidates for T_1 from the reflected intensity distribution are estimated from the average of S_{echo} at the onset delay τ_{on} (4050 μs). Figure 4(e) shows candidates estimated from the average of S_{echo} 10 μs after τ_{on} (4060 μs). Finally, the delay time T_1 is determined as 4100 μs by estimation of the correspondence between reflected intensities of the two averages of S_{echo} at $\tau=4050$ μs and $\tau=4060$ μs.

T_2 was determined by adding T_1 to the estimated delay separation between the first and second objects. Further delay times T_i were determined by iterating these procedures. In the case of many closely spaced objects, S_{echo} at the current delay τ includes interference patterns dependent on the delay separations between these objects. Therefore, it was necessary to remove interference patterns dependent on the delay separations estimated at the previous delay τ. In order to exclude these interferences, S_{ref} at the current delay was computed from the reflected intensity distribution determined at the previous delay. The interference pattern dependent upon the next delay separation was computed by subtracting S_{ref} from S_{echo}, as S_{ref} from closely spaced objects corresponds to the summation of interference patterns dependent on delay separations between these objects (see the Appendix). This subtracted spectrum is labeled S_{sub}.

The cepstrum was computed from S_{sub} by inverse Fourier transform in order to estimate the delay separations for these objects. Since the frequency range was 20 to 120 kHz in this model, the spectrum, which has constant amplitude along the frequency axis, was transformed into the cepstrum, with sidelobes except at $t=0$ through deconvolution of cosine and sine functions. In order to eliminate the sidelobes, these cosine and sine functions were modified into W_{cos} and W_{sin}, respectively (Matsuo *et al.*, 2001).

$$W_{\cos}(f_n,t)=\cos(2\pi t f_n)-\left\{\sum_f \frac{\cos(2\pi t f)}{\mathrm{N}}\right\}, \tag{2}$$

$$W_{\sin}(f_n,t)=\sin(2\pi t f_n)-\left\{\sum_f \frac{\sin(2\pi t f)}{\mathrm{N}}\right\}, \tag{3}$$

where f_n is the center frequency of the cochlear filter and N is the number of cochlear filters. Using the deconvolution of W_{cos} and W_{sin}, S_{sub} was transformed into the cepstra, $C_{cos}(t,\tau)$ and $C_{sin}(t,\tau)$, respectively. The $C(t,\tau)$ were computed using the equation

$$C(t,\tau)=\sqrt{(C_{\cos}(t,\tau))^2+(C_{\sin}(t,\tau))^2}. \tag{4}$$

Determination of the next delay separation from S_{sub} requires estimation of the intensity of the periodic component in the cepstrum C. Since this intensity is represented by $2r_1r_2$ ($r_1 \geqslant r_2$) (see the Appendix), $2r_1r_2$ was computed by extracting the peak of the cepstrum C_{peak}. This value is dependent on both r_1 and r_2. Estimation of the appearance of interference independent of each reflectivity requires the estimation of the correspondence between the reflectivity of two objects. Here, we used the value of C_{peak} divided by C_0 at $t=0$ as the criterion for estimation of appearance of the interference. In the case of two objects, this value corresponds to

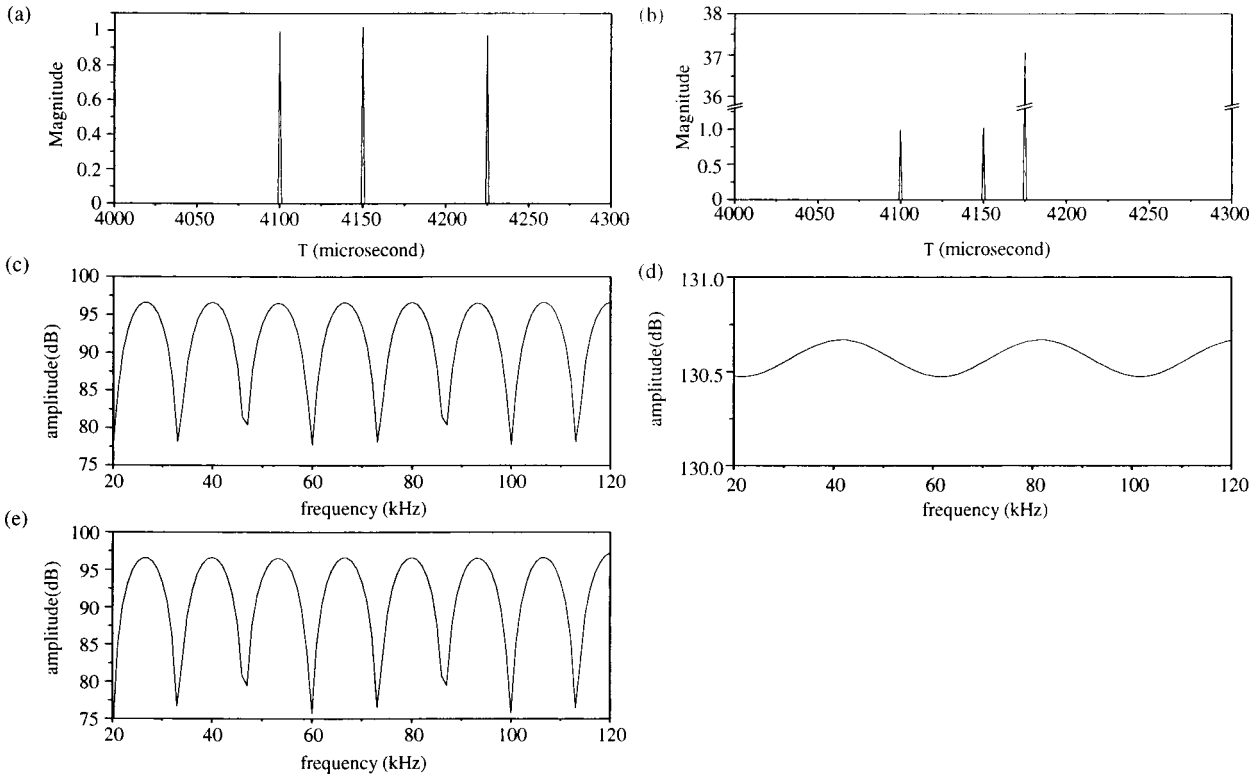


FIG. 5. Determination of the temporal sequence of the delay separations. To find T_3, two candidates in the reflected intensity distribution are estimated from the delay separation determined from S_{echo} at delay τ. (a) Estimated delay separation corresponds to that between the second and third objects. (b) Estimated delay separation corresponds to that between the first and third objects. (c) S_{echo} at delay time τ. (d) S_{ref} computed from the candidate from the reflected intensity distribution in (a). (e) S_{ref} computed from the candidate in the reflected intensity distribution in (b).

$2r_1r_2/(r_1^2+r_2^2)$. When this value was more than 0.385, corresponding to almost 3.5 dB of the difference between the peak and trough amplitudes in the logarithm spectrum, the delay separation was estimated from S_{sub} (see the Appendix).

The determination of each T_i for three or more closely spaced objects requires the determination of the temporal sequence of delay separations. In the process of estimating each delay time T_i, there arise several candidates from the reflected intensity distributions. In the determination of T_3, the estimated delay separation either corresponds to the delay separation between the first and third objects or that between the second and third objects. Here, it was therefore necessary to estimate two candidates for T_3. For each, the reflectivity r_3 was computed using the value of C_{peak}, which was in agreement with the values calculated by multiplying values from the windowed reflected intensity distribution at the delay times found for these two objects. Figures 5(a) and (b) show two candidates from the reflected intensity distributions. In order to determine T_3, S_{ref} at the current delay τ was computed from each candidate of the reflected intensity distribution [Figs. 5(d), (e)]. When the estimated reflected intensity distribution corresponds to the actual reflected intensity, S_{ref} corresponds to S_{echo} at each delay τ. Therefore, the reflected intensity distribution was determined by comparing S_{echo} with S_{ref} at this delay τ. T_i for the second and third objects was determined to be 4150 and 4225 μs, respectively.

III. RESULTS

A. Four closely spaced objects with the same reflectivity

As an illustration of the case for an arbitrary number of objects, we considered a situation with four objects having the same reflectivity. T_i was 4100, 4150, 4190, and 4250 μs, and the reflectivity for these objects was 1. The waveforms, spectrogram, and range–frequency pattern of the echoes returning from these objects are displayed in Figs. 6(a), (b), and (c). T_i for the four objects was determined to be 4100, 4150, 4190, and 4250 μs, and their reflectivities were determined to be 1.000, 1.006, 1.027, and 0.977, respectively, from the reflected intensity distribution in Fig. 6(d). This is in very good agreement with the actual delay times of the objects, while the reflectivities are in reasonable agreement.

B. Three closely spaced objects with different reflectivity

We considered the case of three objects with different reflectivities. T_i for the objects was 4100, 4150, and 4225 μs and their reflectivities were 1.000, 0.100, and 1.000, respectively. The waveforms, spectrogram, and the range–frequency pattern of the echoes from the three objects are displayed in Figs. 7(a), (b), and (c). T_i for the objects were determined as 4100, 4150, and 4225 μs and the reflectivities were found to be 1.001, 0.102, and 0.979, respectively, from the reflected intensity distribution in Fig. 7(d). The calculated

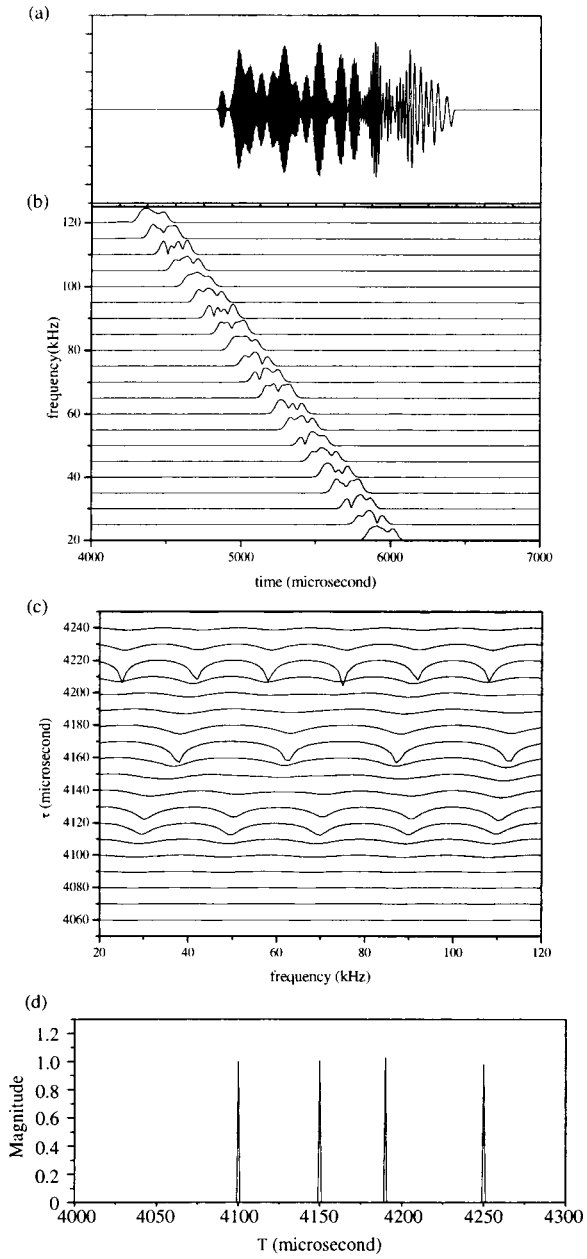


FIG. 6. Four closely spaced objects with the same reflectivity. The delay times are 4100, 4150, 4190, and 4250 μs and the reflectivities are 1. (a) Echo waveform. (b) Time–frequency pattern of echoes. (c) Range–frequency pattern of echoes. (d) Determined reflected intensity distribution.

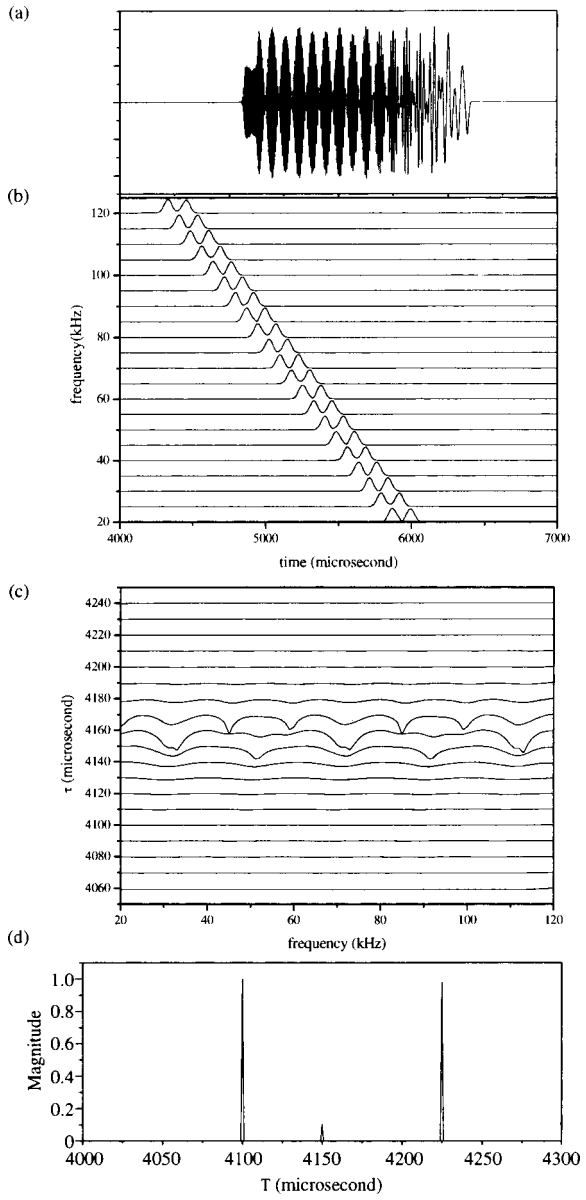


FIG. 7. Three closely spaced objects with different reflectivity. The delay times are 4100, 4150, and 4225 μs, with reflectivities of 1, 0.1, and 1, respectively. (a) Echo waveform. (b) Time–frequency pattern of echoes. (c) Range–frequency pattern of echoes. (d) Determined reflected intensity distribution.

delay times are in very good agreement with the actual delay times, and reflectivities were determined fairly accurately with this model.

C. Resolution of multiple closely spaced objects

In order to evaluate the model's ability to resolve very closely spaced objects, we considered a situation in which the delay separation of echoes from the nearest and farthest objects was 30 μs. T_i was 4100, 4120, and 4130 μs and the reflectivity of all three objects was 1.000. The waveforms, spectrogram, and the range–frequency pattern of the echoes returning from the three objects are displayed in Figs. 8(a), (b), and (c). The model determined the delay times to be 4101, 4121, and 4131 μs with reflectivities of 1.172, 0.995, and 0.894, respectively, from the reflected intensity distribution in Fig. 8(d). Although there are some discrepancies between the actual and calculated reflectivities, it was possible to determine the delay times for the objects with an accuracy of 1 μs.

D. Robustness in a noisy environment

Echolocating bats can reliably discriminate multiple objects in a noisy environment (Griffin, 1958). To investigate the robustness of this model, we added broadband white random noise to the echoes, following Matsuo *et al.* (2001). The noise was characterized by the ratio of the root-mean-square (rms) amplitude of the echo at a single point and the rms amplitude of the white noise during a control interval preceding emission.

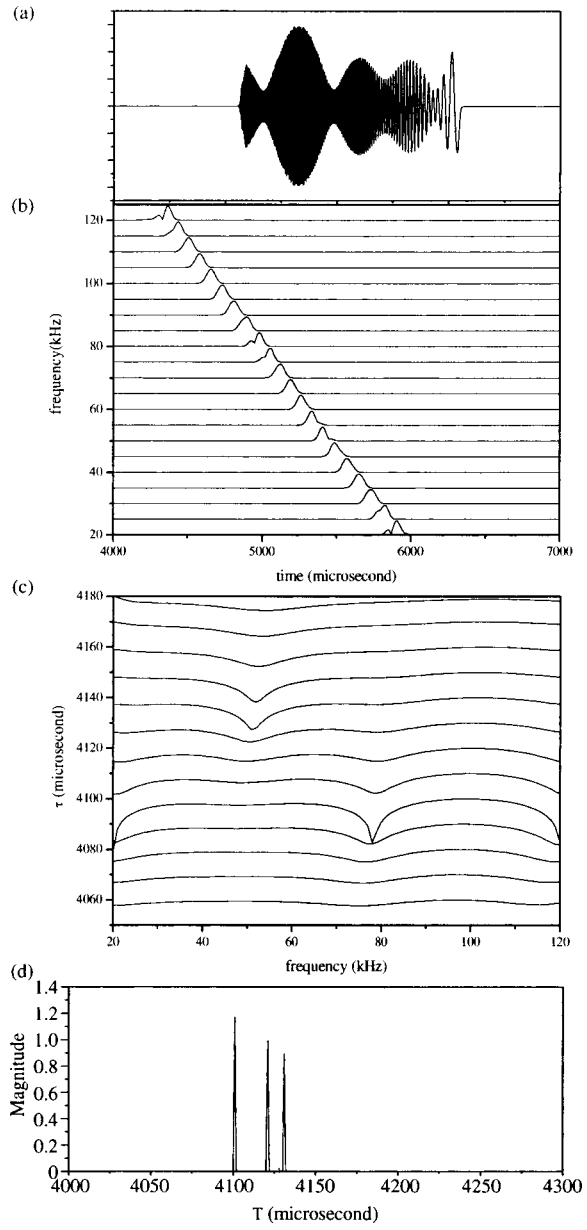


FIG. 8. Resolution of many closely spaced objects. Delay times are 4100, 4120, and 4130 μs, and the reflectivities are 1. (a) Echo waveform. (b) Time–frequency pattern of echoes. (c) Range–frequency pattern of echoes. (d) Determined reflected intensity distribution.

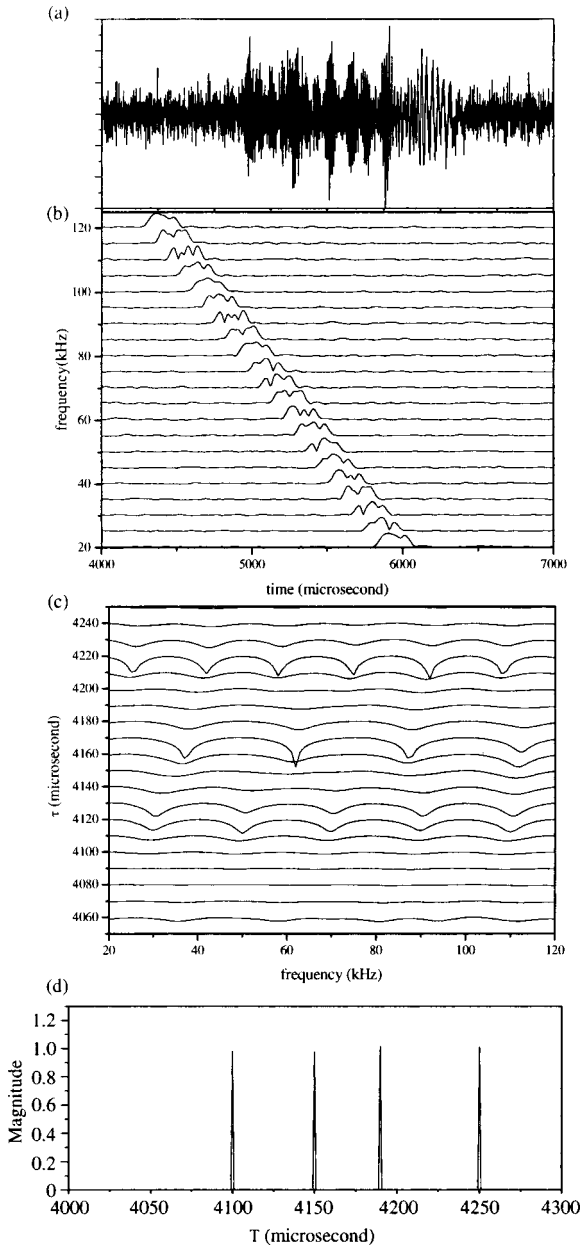


FIG. 9. Robustness in a noisy environment. The delay times are 4100, 4150, 4190, and 4250 μs, and all objects have reflectivity of 1. (a) Echo waveform. (b) Time–frequency pattern of echoes. (c) Range–frequency pattern of echoes. (d) Determined reflected intensity distribution.

In this study, T_i for four objects was 4100, 4150, 4190, and 4250 μs, and the reflectivity of all the objects was 1.000. The rms amplitude signal-to-noise ratio for a single object whose reflectivity corresponds to 1 was 0 dB. The waveforms, spectrogram, and range–frequency pattern of the echoes returning from the four objects are displayed in Figs. 9(a), (b), and (c). T_i calculated for the four objects was 4100, 4150, 4190, and 4250 μs, and their reflectivities were determined to be 0.981, 0.975, 1.015, and 1.009, respectively, from the reflected intensity distribution in Fig. 9(d). The calculated delay times agree with the actual times, while these reflectivities were determined with reasonable accuracy.

IV. DISCUSSION AND CONCLUSION

In the case of many closely spaced objects, the echoes interfere with one another. In order to determine each T_i, it is necessary to solve two problems: how to accurately determine T_1, and how to determine the temporal sequence of the delay separations.

In relation to the first problem, experimental evidence indicates that FM bats are capable of locating an object with better than microsecond accuracy (Menne *et al.*, 1989; Moss and Schnitzler, 1989; Simmons, 1979; Simmons *et al.*, 1990). In contrast, it has been found in physiological studies that responses in the neurons of the inferior colliculus of the FM bat, *Tadarida brasiliensis mexican*, have temporal precision of hundreds of microseconds (Pollak *et al.*, 1977). Other studies have found that some neurons in the nucleus of the lateral lemniscus of the FM bats, *Eptesicus fuscus*, have latency variability as small as 20 to 50 μs (Covey and Casseday, 1991; Haplea *et al.*, 1994). It seems that it is difficult to

determine T_1 from neurons having these response times.

It is possible to determine T_1 using the cross-correlation function between spectrograms of both the emission and echoes. However, it is inevitable that this will produce an estimated error of several tens of microseconds, because the integration time of the cochlear filters is several hundred microseconds (Matsuo *et al.*, 2001). In contrast, here we used Gaussian chirplets to transform the echoes into spectra, so that they almost agreed with the S_{ref} computed from the reflected intensity distribution and the temporal emission pattern W_{emi} as shown in Fig. 3. This shows that the echo spectra are a function of not only the delay time but also the reflectivity for each object. In order to determine T_1 and r_1, echo spectra at least two different delay times are required. In this model, T_1 was determined from the echo spectrum at the onset delay τ_{on} and the spectrum 10 μs later. This indicates that the estimated T_1 becomes independent of the integration time of the cochlear filters. In the case of three objects located within 30 μs of delay separation, T_1 can be determined from two echo spectra with an accuracy of a microsecond as shown in Fig. 8(d). Even in a noisy environment having a signal-to-noise ratio of 0 dB, T_1 can still be accurately determined as shown in Fig. 9(d). These results suggest that this proposed model might explain the discrepancy between the measured physiological low temporal precision and the behavioral high temporal precision.

The next problem is to accurately determine the temporal sequence of delay separations. This requires the extraction of temporal changes in the interference patterns from the echo spectra. However, temporal changes in the interference patterns cannot be extracted from the outputs of cochlear filters with constant carrier frequency [Fig. 2(c)]. Here, we used Gaussian chirplets having carrier frequency in agreement with the sweep rate of the emission. In the case of three objects located within 30 μs of delay separation, Fig. 8(c) shows that the interference pattern originating from echoes from the first and second objects can be discriminated from that between echoes from the second and third objects. Thus, the temporal sequence of delay separations among these closely spaced objects can be extracted.

In these procedures, we first introduced the concept of extracting the interference of echoes as the criterion for determination of the temporal sequence of delay separations. Since the echo spectra are varied with delay τ by convolution with Gaussian chirplets, we used the value of C_{peak}/C_0 as the criterion. In the case of reflectivity r_1 and r_2 for two objects, this criterion is $2r_1r_2/(r_1^2+r_2^2)$. This takes its maximal value of 1 when r_1 and r_2 are equal. In this model, we used 0.385 as a measurable threshold, which corresponds almost to 3.5 dB of the difference between the peak and trough amplitudes of the logarithm spectrum (see the Appendix). This is similar to the threshold for detecting the periodicity of a spectrum for humans and budgerigars (Amagai *et al.*, 1999).

Finally, we determined the temporal sequence of delay separations for three or more closely spaced objects. T_2 was uniquely determined from the first appearance of delay separation. In the case of a second delay separation, it cannot be determined whether this corresponds to the separation between the first and third objects or the second and third objects, so there are two candidates for T_3. In each case, r_3 was obtained from C_{peak} and the candidate T_3. S_{ref} was then computed from the candidates for T_3 and the reflectivity r_3. In this model, T_3 was then determined by measuring the correspondence between the echo spectra and S_{ref} computed for each candidate. By iterating these procedures, each subsequent delay time T_i may be determined in order. It was successfully demonstrated that each T_i for three or more closely spaced objects can be determined to within 1-μs accuracy. This is possible even when the three objects are all located within 30 μs of delay separation, the reflectivity of the objects is different, and when the objects are in a noisy environment with a 0-dB signal to noise ratio, as shown in Figs. 6, 7, 8, and 9.

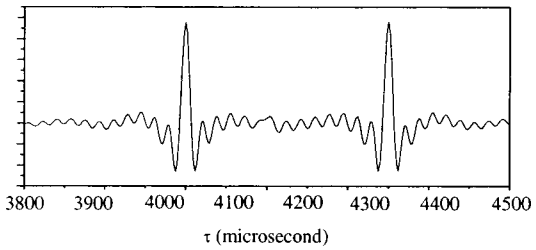


FIG. 10. Cross-correlation function between emitted and reflected waveforms, for three objects. The delay times are 4100, 4150, and 4225 μs, with reflectivities of 1, 0.1, and 1, respectively. T_2 cannot be determined from this cross-correlation function.

It should be noted that there is a large discrepancy between the delay times calculated using this method and those calculated using the cross-correlation function between the waveforms of emission and reflection. The delay times found using the latter method are determined from the peak times of the cross-correlation function. Figure 10 shows the cross-correlation function in the case of three objects with different reflectivity. The delay times for these objects are 4100, 4150, and 4225 μs, and reflectivities are 1.000, 0.100, and 1.000, respectively. The small peaks, labeled ''sidelobe peaks,'' appear around delay times for the first and third objects because the emitted frequency ranges only from 20 to 120 kHz. The value of the cross-correlation function at the delay time for the second object is smaller than the sidelobe peak. This means that it is difficult to determine delay times for closely spaced objects with different reflectivity using that method.

To locate multiple objects in 3D space using echoes from a single emission requires directional information as well as the distance of each object. We previously proposed the binaural model to determine the location of two closely spaced objects in 3D space by estimating the directional information from the echo spectrum. The present model can be easily extended to a binaural model to determine the location of multiple closely spaced objects in 3D space.

APPENDIX: SPECTRA COMPUTED FROM THE REFLECTED INTENSITY DISTRIBUTION

The delay times for two objects are T_1 and T_2, and the reflectivities of two objects are r_1 and r_2, respectively. The interference spectrum computed from the positions of the two objects is represented by the equation

$$S(f)=r_1^2+r_2^2+2r_1r_2\cos(2\pi f(T_2-T_1)).$$

The peak and trough values of this spectrum correspond to $r_1^2+r_2^2+2r_1r_2$ and $r_1^2+r_2^2-2r_1r_2$, respectively. When the peak-to-trough logarithmic amplitude is 3.5 dB, the ratio of r_2 and r_1 (given by r_2/r_1) corresponds to 0.2 and the value of $2r_1r_2/(r_1^2+r_2^2)$ corresponds to 0.385.

The delay times for three objects are T_1, T_2, and T_3, and their reflectivity is given by r_1, r_2, and r_3, respectively. The interference spectrum computed from positions of the three objects is represented by the equation

$$\begin{aligned}S(f)=&r_1^2+r_2^2+r_3^2+2r_1r_2\cos(2\pi f(T_2-T_1))\\&+2r_2r_3\cos(2\pi f(T_3-T_2))+2r_3r_1\\&\times\cos(2\pi f(T_3-T_1)).\end{aligned}$$

Altes, R. A. (**1976**). ''Sonar for generalized target description and its similarity to animal echolocation systems,'' J. Acoust. Soc. Am. **59**, 97–106.

Amagai, S., Dooling, R. J., Shamma, S., Kidd, T. L., and Lohr, B. (**1999**). ''Detection of modulation in spectral envelopes and linear-rippled noises by budgerigars (*Melopsittacus undulates*),'' J. Acoust. Soc. Am. **105**, 2029–2035.

Covey, E., and Casseday, J. H. (**1991**). ''The monaural nuclei of the lateral lemniscus in an echolocating bat: Parallel pathways for analyzing temporal features of sound,'' J. Neurosci. **11**, 3456–3470.

Fenton, M. B., and Bell, G. P. (**1981**). ''Recognition of species of insectivorous bats by their echolocation calls,'' J. Mammal. **62**, 233–243.

Griffin, D. R. (**1958**). *Listening in the Dark* (Yale University Press, New Haven CT; reprinted by Cornell University Press, Ithaca, NY, 1986).

Haplea, S., Covey, E., and Casseday, J. H. (**1994**). ''Frequency tuning and response latencies at three levels in the brainstem of the echolocating bat, *Eptesicus fuscus*,'' J. Comp. Physiol., A **174**, 671–683.

Mann, S., and Haykin, S. (**1992**). '' 'Chirplets' and 'wavelets': Novel time-frequency methods,'' Electron. Lett. **28**, 114–116.

Mann, S., and Haykin, S. (**1995**). ''The chirplet transform: physical considerations,'' IEEE Trans. Signal Process. **43**, 2745–2761.

Matsuo, I., Tani, J., and Yano, M. (**2001**). ''A model of echolocation of multiple targets in 3D space from a single emission,'' J. Acoust. Soc. Am. **110**, 607–624.

Menne, D., Kaipf, I., Wagner, I., Ostwald, J., and Schnitzler, H. U. (**1989**). ''Range estimation by echolocation in the bat *Eptesicus fuscus*: trading of phase versus time cues,'' J. Acoust. Soc. Am. **85**, 2642–2650.

Moss, C. F., and Schnitzler, H. U. (**1989**). ''Accuracy of target ranging in echolocating bats: Acoustic information processing,'' J. Comp. Physiol., A **165**, 383–393.

Neretti, N., Sanderson, M. I., Intrator, N., and Simmons, J. A. (**2003**). ''Time-frequency model for echo-delay resolution in wideband biosonar,'' J. Acoust. Soc. Am. **113**, 2137–2145.

Pollak, G. D., Marsh, D. S., Bodenhamer, R., and Souther, A. (**1977**). ''Characteristics of phasic-on neurons in the inferior colliculus of unanesthetized bats with observations related to mechanisms for echo ranging,'' J. Neurophysiol. **40**, 926–941.

Saillant, P. A., Simmons, J. A., Dear, S. P., and McMullen, T. A. (**1993**). ''A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats: The spectrogram correlation and transformation receiver,'' J. Acoust. Soc. Am. **94**, 2691–2712.

Simmons, J. A. (**1979**). ''Perception of echo phase information in bat sonar,'' Science **207**, 1336–1338.

Simmons, J. A., Ferragamo, M. J., and Moss, C. F. (**1998**). ''Echo-delay resolution in sonar images of the big brown bat, *Eptesicus fuscus*,'' Proc. Natl. Acad. Sci. U.S.A. **95**, 12647–12652.

Simmons, J. A., Ferragamo, M. J., Moss, C. F., Stevenson, S. B., and Altes, R. A. (**1990**). ''Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: The shape of target images in echolocation,'' J. Comp. Physiol., A **167**, 589–616.

Simmons, J. A., Ferragamo, M. J., Saillant, P. A., Haresign, T., Wotton, J. M., Dear, S. P., and Lee, D. N. (**1995**). ''Auditory dimensions of acoustic images in echolocation,'' in *Hearing by Bats*, edited by A. N. Popper and R. R. Fay (Springer, New York), pp. 146–190.

Simmons, J. A., Saillant, P. A., Ferragamo, M. J., Haresign, T., Dear, S. P., Fritz, J., and McMullen, T. A. (**1996**). ''Auditory computations for biosonar target imaging in bats,'' in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Springer, New York), pp. 401–468.

Erratum: "The dependency of timbre on fundamental frequency" [J. Acoust. Soc. Am. 114, 2946–2957 (2003)]

Jeremy Marozeau, Alain de Cheveigné, Stephen McAdams, and Suzanne Winsberg
Institut de Recherche et Coordination Acoustique/Musique (Ircam-CNRS), 1, place Igor Stravinsky, F-75004 Paris, France

(Received 26 November 2003; accepted for publication 28 November 2003)

[DOI: 10.1121/1.1642764]

PACS numbers: 43.75.Cd, 43.66.Jh, 43.66.Hg, 43.10.Vx [ADP]

The citation ''Ueda and Nimmo-Smith (1987)'' and corresponding reference are incorrect. The correct reference is: Ueda, K., and Ohgushi, K. (1987). ''Perceptual components of pitch: Spatial representation using a multidimensional scaling technique,'' J. Acoust. Soc. Am. **82**, 1193–1200.

